Artificial Intelligence in Higher Education: A Research Pathway with ChatGPT for Learning Design, Feedback, and Professional Development

Maila Pentucci*, Manuela Fabbri**, Chiara Laici***°

Abstract

This paper explores a research pathway that leverages an AI-based conversational tool, ChatGPT - OpenAI, to enhance essential competencies in future teachers and educators, with a focus on self-reflection and feedback literacy. Conducted within two pedagogical courses, the activity involved peer feedback on didactic design tasks, fostering students' agency and metacognitive reflection. By using ChatGPT as both a design and feedback agent, students evaluated its effectiveness, strengths, and limitations. Reflective questionnaires allowed them to assess the tool's potential integration into their future professional practices, addressing the broader applicability of AI in educational contexts.

Key words: feedback; learning design; AI; ChatGPT; professional development.

* Università "d'Annunzio" di Chieti-Pescara. E-mail: maila.pentucci@unich.it (corresp. author).
** Università degli Studi di Bologna. E-mail: m.fabbri@unibo.it.
***Università degli Studi di Macerata. E-mail: chiara.laici@unimc.it.
° This paper is the result of a collaborative effort among the authors. The roles are attributed as follows: **Maila Pentucci:** development of the research framework; data collection and analysis; writing of Chapters **1** ("Introduction"), **2** ("Background"), and **4.2** ("Strengths and Weaknesses of ChatGPT's Feedback"). **Manuela Fabbri:** Research implementation; data collection and analysis; writing of Chapters **3** ("Description of the Context and Research Framework") and **4.3** ("ChatGPT and Professional Development"). **Chiara Laici:** Data analysis; critical review of the research framework; writing of Chapters **4.1** ("Strengths and Weaknesses of Learning Design") and **5** ("Conclusions"). Moreover, **ChatGPT (OpenAI):** Performed statistical processing of certain data and generated the illustrative graphs.

1. Introduction

This paper presents a research pathway, using an Artificial Intelligence-based conversational tool, to enhance and support key competencies in preparing future teachers and educators, particularly self-reflection (van Velzen, 2015) and feedback literacy (Nieminen and Carless, 2023). Specifically, we used ChatGPT as an interactive tool in a peer feedback and review activity with students (Pentucci and Laici, 2023). The activity focused on a didactic design task assigned within university courses in the pedagogical-didactic field at two different universities.

From an ecosystemic perspective, we deemed it appropriate to integrate the tool in a way that emphasises both agency and proactivity (Rivoltella and Rossi, 2019). Our two main objectives were: 1) to replicate the same task assigned to students, allowing for a comparative analysis of outputs and fostering metacognitive reflection on practice; and 2) to provide feedback on students' work, enabling them to evaluate the feedback and decide how, or if, to revise their tasks.

This hands-on experience also enabled students to reflect on the application of Artificial Intelligence in educational settings, assessing both the strengths and limitations of the ChatGPT tool when tested in such interactive processes.

Through self-reflection questionnaires, students were able to articulate their perceptions of this usage and share their thoughts on the potential integration of the tool within their professional practices. In this context, we guided them through a process of self-examination focused on both their present training in educational fields and their future roles as teachers and educators.

The research questions we aim to address through an analysis of the students' responses are as follows:

a) What strengths and limitations were observed when using the tool as both a design agent and a feedback agent?
b) Can the tool support, initiate, or facilitate feedback and reflection processes in teacher and educator training?
c) Based on students' direct experience with the tool, what assumptions can be made about its application within the professional practices required in the teaching profession?

2. Background

Pre-trained generative systems based on artificial intelligence, with their ability to interpret and generate text interactively, have become common tools in everyday life (Hwang and Chang, 2021; Perera and Lankathilaka, 2023).

Their interpretive algorithms and rapid data processing give them a conversational quality that makes them resemble human agents (Deng et al., 2024).

This feature enables interaction via natural language instead of coding (Liao et al., 2023).

The most significant educational studies (McNamara et al., 2023; Prananta et al., 2023) addressing future challenges and developments emphasise proactive applications of AI (Panciroli and Rivoltella, 2023; Følstad et al., 2021), positioning AI as a partner, facilitator, or co-author within educational frameworks. This approach aligns with the concept of an AI culture (Elliot, 2021), where interactions between human and non-human agents become pivotal.

Indeed, the agentive and relational characteristics previously highlighted position AI-based generative systems as, in many respects, subjects rather than mere objects within integrated educational ecosystems (Gupta et al., 2024). These systems can transform other agents and, reciprocally, be transformed by them through targeted actions based on a deep understanding of their logic and functionality, such as prompting and training capacity (Wan and Chen, 2024; Walter, 2024).

This shift requires a rethinking of teaching practices in light of these new interactive partners: on the one hand, there is a need to understand new paradigms based on dataisation (Wiliamson et al., 2023) and digital plenitude (Bolter, 2019), while on the other, it is essential to explore the roles and functions that AIs might fulfill in educational settings, particularly in higher education. By leveraging conversational capabilities, these tools can serve as feedback agents (Gratani et al., 2023; Pang et al., 2024), encouraging the development of students' feedback literacy across its various dimensions (Carless and Boud, 2018): learning to appreciate and interpret the feedback received, managing emotions and anxieties associated with evaluation, acquiring the skills to process and provide constructive feedback, and formulating questions that foster self-awareness and critical comparison. Above all, these tools support students in using feedback in a generative manner (Rossi et al., 2018), enabling them to restructure and improve their performance and learning processes.

Exploring the potential of AI in developing future teachers' professional competencies, such as learning design (Rossi and Pentucci, 2021) and the ability to provide feedback with transformative intent (Winstone and Carless, 2019; Laici, 2021; Sansone, Bortolotti and Fabbri 2021), must necessarily begin with co-experimentation and co-research practices alongside students in training.

This approach aims to surface the knowledge underpinning these practices and underscores their importance in consolidating a reflective stance (Pentucci, 2018; Sansone, Fabbri and Bortolotti, 2023).

3. Description of the Context and Research Framework

The research was carried out in the context of two university degree programmes in the pedagogical field. The participants were 62 students, equally divided into two groups: 31 third year students of the Degree in Pedagogical Sciences at the University "d'Annunzio" of Chieti (Italy) and 31 second year students of the Degree in Mathematics at the University of Bologna (Italy).

The research framework was divided into several phases. These are outlined in Fig 1.

Phases of the exploratory study	Description
1. Design of the educational pathway	Small group design of an educational or disciplinary pathway focused on a specific target of choice
2. Generation of an educational path by ChatGPT	Design of the same task by ChatGPT, following the same instructions
3. Analysis of the ChatGPT product: individual reflection	Fill in a questionnaire (a) in order to a) evaluate the product generated by ChatGPT b) compare the product generated by ChatGPT with that of the group
4. Analysis of the ChatGPT product: group reflection	Fill in a 'comparison sheet' provided by the teachers in order to a) evaluate the product generated by ChatGPT b) compare the product generated by ChatGPT with that of the group c) identify the strengths and weaknesses of the ChatGPT generated product
5. ChatGPT's analysis of the pedagogical design created by the students	ChatGPT corrects and evaluates the task created by each group based on a prompt that provides accurate feedback in terms of both strengths and areas for improvement
6. Group reflection and possible revision of their own text	Group reflection on the feedback received from ChatGPT by filling in a 'structured sheet' provided by the teachers
7. IIndividual reflection and possible revision of their own text	Completion of a questionnaire (b) to metacognitively express reflections and activate potential transformative processes
8. Reorganization of knowledge	Joint lecture by the teachers on AI and ChatGPT in the educational context

Fig. 1 - Phases of the exploratory study

The students worked both individually and in small groups to design an educational pathway on a specific topic of their choice, which was later compared with a similar design generated by ChatGPT. Through a series of questionnaires and reflective activities students were asked to evaluate the quality of ChatGPT's work, compare it with their own, and reflect on the potential future uses of such tools in educational contexts. Moreover, ChatGPT was utilised to evaluate the students' projects, providing feedback that the groups analysed and discussed to refine their work. The findings from this process were used to explore how generative AI can be integrated into the

teaching and learning processes, aiming to enhance both teacher and student agency in educational settings.

The two questionnaires were extensive and detailed and provided significant data on various aspects. They were given to the students and focused on two dimensions.

In relation to the present, students reflected on the potential and limitations of ChatGPT in the context of their educational experience. In particular, they discussed how AI could facilitate lesson planning, feedback and critical reflection, while also identifying challenges such as the need for constant human supervision; looking to the future, students reflected on the future use of ChatGPT in their professional practice. They were asked to consider whether they would use this tool in their future work as educators/teachers, and in what contexts it might be useful, both as an aid to design and assessment, and as a tool for their students to engage in self-assessment and autonomous learning.

Specifically, the first questionnaire (a), completed by 61 students (44 F and 16 M; mean age 23.2 years), asked participants to evaluate the product generated by ChatGPT and to compare it with that of their own group. The second questionnaire, completed by 57 students (42 F and 15 M; mean age 23.4 years), focused on both the feedback received from the ChatBot in relation to their own task and the potential future uses of ChatGPT, both as an educational professional and as a tool for students.

Questionnaire number 1

Answers:
1. What are the strengths of the task created by the chatbot? (open answer)
2. What are the weaknesses of the task created by the Chatbot? (open answer)

Questionnaire number 2

Answers:
1. What are the weaknesses of the feedback (suggestions, modifications, etc.) received from the Chatbot regarding your group's task? (open answer)
2. What are the strengths of the feedback (suggestions, modifications, etc.) received from the Chatbot
3. Would you use ChatGPT in an educational context as a tool for teachers/educators? (closed answer)
4. For which activities would you use it? (open answer)
5. Would you let your students/learners use ChatGPT? (closed answer)
6. For which activities would you let them use it? (open answer)

Fig. 2 – Questionnaire's answers

The open-ended responses were analysed using Braun and Clarke's (2019) model: Reflexive Thematic Analysis. This analysis involves an inductive approach, as tags are identified from the qualitative and reflexive interpretation of the data itself. This approach seeks to bring out the latent meanings, ideas and conceptualisations underlying the data. The codings were reviewed and

agreed upon by the three authors (Braun and Clarke, 2012). A frequency analysis was then performed on the tags, following the principles of qualitative content analysis (Krippendorff, 2019).

This analysis allowed us to identify and quantify recurring themes and concepts "in order to organise the story into a coherent and internally consistent account" (Braun and Clarke, 2006, p. 22).

We also conducted an analysis of co-occurrences between tags by representing them through heatmaps that highlighted certain patterns of data concentration.

4. Results and discussion

4.1 Strengths and weaknesses of learning design

The first questionnaire asked the students to analyse the strengths and weaknesses of the learning design made by ChatGPT, comparing it with that made by their own group.

The students therefore approached the task produced by ChatGPT not from a simple evaluative or revision perspective, but in a posture of comparison and activation of an inner feedback (Nicol, 2020; Nicol and McCallum, 2022; Tam, 2024). They were able to reflect and were able to mirror their own cognitive mechanisms in the artefact. They also reflected on the strategies activated and the difficulties detected and faced in their own design actions.
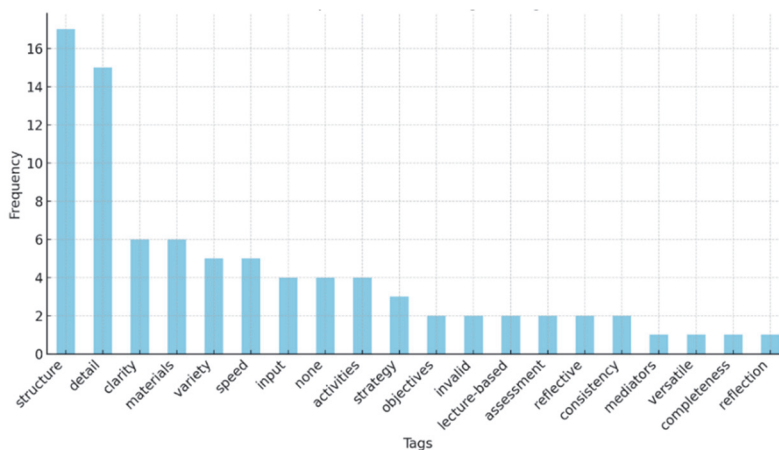

Fig. 3 - *Distribution of absolute frequencies of the tags related to the strengths of the learning design realised by the Chatbot*

With respect to strengths, as evidenced by Figure no. 3, the most frequent tags that can be associated with the 61 students' reflections are *structure* (N = 17) and *detail* (N = 15), followed by fewer tags associated with *clarity*, *materials* variety and *speed*.

The analysis of the co-occurrences of the tags within the same answers also shows that the tag *structure* recurs 5 times with *clarity*, 4 times with *detail* and 4 times with *speed*. 4 times with *detail* and 3 times with *variety*.

The learning design implemented by the Chatbot is perceived by the students as well organised and structured. A clear and logical organisation of the activities was provided, effectively dividing the content and thus facilitating orientation through the various work steps. The combination of the tags *structure* and *clarity* shows that the Chatbot provides adequate support in breaking down the content into clear, well-defined and easy-to-follow units (*The division of activities is clear and well-defined, which makes it easy to follow the whole learning path; I appreciate the structured sequence of activities, which helps to maintain order during the lesson*).

Attention to detail further enhances the quality of the design realised by the Chatbot, making it easier to understand the activities and learning objectives. According to the respondents, clarity and precision provide an easy-to-follow learning path, allowing a clearer idea of how to proceed.

In fact, the Chatbot seems to not only provide a general structure, but also includes a significant level of depth, ensuring that teachers have much of the information they need to conduct the teaching activities effectively and completely (*The activity is well detailed, each step is described precisely and completely; The details provided by the chatbot helped me to clearly understand what to do at all times*).

The most significant co-occurrences *structure-clarity* and *structure-detail*, thus highlight that according to the students, Chabot is able to work positively on a design framework that has the characteristics of clarity and accuracy, thus generating a complete artefact that responds to the students' idea of learning design.

This tendency is not surprising: in fact, novices tend to favour, in the design of the microsession, more talking artefacts, with a higher level of detail and description (Bonaiuti et al., 2017), which can support, guide and orient more firmly.

The transition to an open and fluid design is realised when we take possession of more solid professional reflexive postures thanks to experience and the awareness that, in order to govern the unexpected, we need a design that orients but does not cage, that encompasses deviation, redundancy, scaffolding to immediate decision-making and regulation in action (Capolla et al., 2024; Pentucci, Rossi and Capolla, 2023).

A further reflection can be made regarding the materials suggested and the variety of proposals made by the IA.

Both future teachers and educators pointed out that the Chatbot used a good variety of resources and tools, the materials were perceived as adequate and functional. Students recognise that the Chatbot provides useful resources consistent with the activities, enhancing the learning experience (*The materials provided are adequate and integrate well with the proposed activities; The presence of resources is a great advantage because it facilitates teaching and makes learning more complete*).

According to the respondents, variety helps to keep students' attention and involvement high, while the materials provide practical and didactic support. Indeed, the adoption of different approaches in the teaching process promotes more dynamic and engaging learning, adaptable to different approaches and educational needs (*The variety of activities proposed keeps students' attention high; The fact that there are different types of activities makes the lesson more dynamic and engaging*).

We can also point out that the speed with which the Chatbot generates tasks and resources is seen as an advantage, allowing teachers who have to manage multiple tasks to save time in preparing lessons. The speed allows them to focus more on other aspects of teaching, such as personalisation/individualisation of learning or interaction with students (*The Chatbot is fast in generating tasks and this saves me a lot of time in preparation; The speed of the design process is really useful, especially when you have many other things to do*).

Overall, the strengths of the design realised with the Chatbot can be summarised in a clear and detailed structure, the variety of activities and the speed of execution. These aspects suggest that the learning design realised by the Chatbot can be a valuable tool to support teachers and educators in their professional activity.

With respect to the weaknesses that students identify in the ChatGPT design, tag analysis (Fig. 4) shows that the most frequent include *generic* (N = 18) and *timing* (N = 13), followed by *lecture-based* (N = 8) and *knowledge-based* (N = 7). Among the most frequent co-occurrences between tags are *generic* and *lecture-based* 3 times, *knowledge-based* and *non-human* 3 times and a series of co-occurrences related to the *timing* tag: *timing* and *lecture-based*, *timing* and *generic*, *timing* and *materials* which occur 2 times each.

The main limitation of the learning design implemented by ChatGPT is that many parts are generic or superficial. Several student responses highlight that although there are many details and a good structure, there is a lack of depth in the contents, with the feeling that the activities and resources are standardised and not sufficiently adapted to the needs of the students or the educational context and therefore almost aseptic. This criticism reflects a desire for greater

specificity and relevance in the didactic proposals (*Some activities seem too generic to me, they would need to be customised for the specific context; There is a lack of personalisation to make the didactic design more suitable for different types of students*).
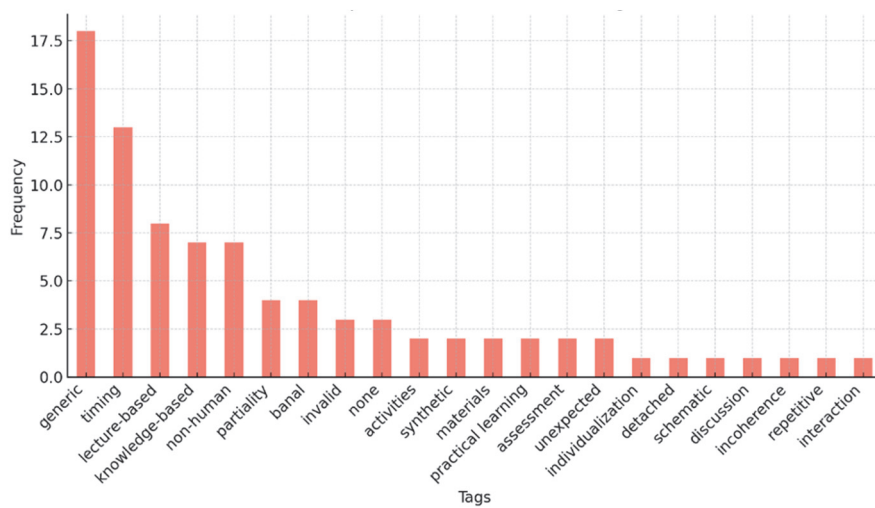

Fig. 4 - Absolute frequency distribution of the tags related to the weaknesses of the learning design realised by the Chatbot

Moreover, many students find that the Chatbot favours a traditional teaching approach, with too much emphasis on lectures and transmission of notions or activities that could be taken from any textbook. This type of approach, although useful in some cases, is seen as limiting it in favour of more active, hands-on, interaction-based learning (*The design relies too much on lectures, without leaving enough opportunities for interaction; This approach seems too traditional to me, it does not actively involve students*). In fact, this aspect is consistent with one of the characteristics of the Large Language System (LLS) and their training: the massive amount of data referring to educational theories and examples of practice includes old and new, traditional and innovative with a clear preponderance of the first type of data.

This confirms one of the aspects relating to AI that has been given particular attention today with regard to the sociological and ethical dimension: AI mirrors and renders reality, both in moral and social questions and in more marginal issues such as the one we are dealing with. Traditional didactics and classical modes of teaching are indeed majority, and AI reflects this trend (Packin and Lev-Aretz, 2018; Coeckelbergh, 2020).

Another reflection related to the limitations of the design implemented by ChatGPT concerns time. According to the students, in fact, the proposed activities are often not well balanced in terms of timing, with some taking too much or too little time with respect to the expected objectives, compromising the effectiveness of the lesson (*The activities seem too long for the time I have available; The time management is not well calibrated with respect to the objectives to be achieved*).

This is an interesting aspect because time is a substantial and highly challenging element in the instructional design process, strongly constraining and at the same time highly connected to the contextualisation of the action and the designer's knowledge of the ecosystem in which one acts (Rossi and Pentucci, 2021). The fact that students realise the AI tool's inability to design time effectively and coherently means that they possess an awareness of this aspect, which significantly connotes the teacher's design competence.

Finally, interesting is the perceived weakness of the design attributed to the lack of empathy and human interaction. The design produced by the Chatbot appears detached, cold and mechanical, failing to replicate the interactive dynamics that are fundamental in an educational context in which communication between teacher and students plays a crucial role (*The Chatbot's approach is very mechanical, it lacks the human element that usually enriches teaching; There seems to be a lack of real interaction, as if the process were too automatic and detached*).

On the whole, the weaknesses of the learning design realised by ChatGPT can be attributed to proposals that are too generic and not adapted to the context or the students, proposals that are based on a predominantly traditional didactics using the frontal lecture, that fail to manage time optimally and that are often mechanical and cold.

The following figure (Fig. 5) summarises the main strengths and weaknesses of the design by thematization.

Strengths	Explanation
Clarity of structure	ChatGPT organises the learning path in a clear and well-structured manner, making it easy to understand the various work steps and objectives.
Attention to detail	The design produced by ChatGPT includes precise and comprehensive details, which are useful to ensure accurate execution of activities.
Richness of materials	ChatGPT provides a variety of supporting teaching materials, facilitating more complete and varied learning.
Variety of activities	The didactical proposals are varied and offer different types of approaches, adapting to different learning modes and keeping motivation high.
Efficiency and speed	ChatGPT's ability to quickly produce well-structured routes is seen as a major time-saving advantage.
Weaknesses	**Explanation**
Lack of customisation	The proposed activities are often perceived as too generic and not adapted to the specific needs of the students or the teaching context.
Lack of human interaction	ChatGPT is perceived as cold and distant, lacking the ability to provide empathic feedback and disregarding students' relational needs.
Traditional approach	Design tends to be based on an overly frontal and transmissive approach, which does not actively involve students in practical or interactive activities.
Ineffective time management	The proposed activities often do not respect the time available, being either too long or too short in relation to the set objectives.
Superficiality in some activities	Although there is variety, some activities are perceived as too simple or banal, without offering enough challenge or inspiration.

Fig. 5 - Main strengths and weaknesses of the learning design

In order to explore the students' overall thinking, co-occurrences emerging between the strengths and weaknesses of the ChatGPT learning design were also examined (Fig 6).
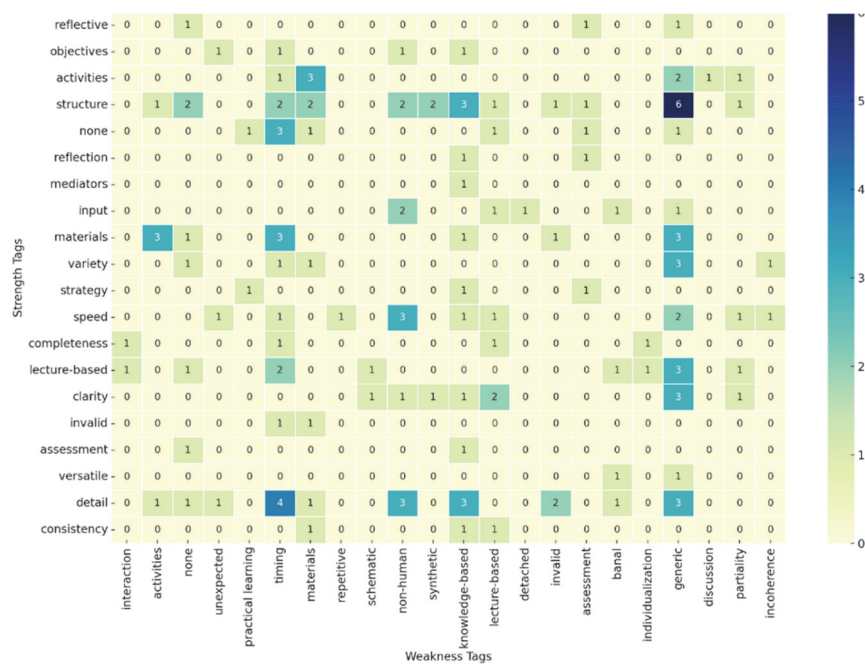
Fig. 6 - Co-occurrences between tags related to weaknesses and strengths in the design implemented by the Chatbot

The strongest relationships between the tags include *structure* and *generic* with 6 co-occurrences and *detail* and *timing* with 4 co-occurrences. The *generic* tag is also often associated with other tags such as *variety*, *lecture-based*, *detail* and *materials*, each with 3 co-occurrences. The *detail* tag is also associated with *knowledge-based* and n*on-human*, each with 3 co-occurrences.

Indeed, the students seem to appreciate the ability to construct a good design structure, but have reservations about the generic nature of the content or the notionism of the proposed design. Similarly, those who appreciate the details and precision of the work, complain about an incorrect management of time (*It looks like the chapter of a normal school book, which is a weakness in my opinion, because it doesn't add anything, in fact the teacher can step aside and put in someone who knows the concepts and follows this chapter to have the same effect*).

Another aspect that emerges is the tendency to humanise the Chatbot. In fact, when commenting on the task, students immerse themselves in the peer review process and tend to talk about ChatGPT as if it were a human agent and designer: this implication leads them to point out, among the weaknesses, certain factors that have been classified as non-human. The non-human tag

links to reflections in which the technical strengths of AI-designed lessons (such as organisation, clarity and precision) and the perceived lack of human warmth, empathy or involvement emerge. Although the students recognise the practical benefits of using AI to design lessons, they perceive how these benefits are overshadowed by the absence of human and personalised elements that would make learning more engaging and emotionally rewarding (*The activity produced by ChatGPT appears schematic and orderly and therefore easily understandable but the weaknesses are that since it is an artificial intelligence, it does not take into account the dynamics that occur within the classroom*).

In summary, the learning design implemented by ChatGPT shows significant potential according to the students, offering a clear structure, attention to detail and speed of execution. However, the generic nature of the proposals, which are often too traditional, the limited empathic interaction and time management weaknesses represent challenges and elements to be considered for design improvement.

4.2 Strengths and Weaknesses of ChatGPT's Feedback

Questionnaire no. 2 focused on feedback. We asked the students what they perceived to be the strengths and weaknesses of the feedback ChatGPT gave to their task.

The coding of responses regarding the strengths of the feedback generated by ChatGPT reveals a clear predominance of the tag *integration*, with 21 occurrences (see Fig. 7). These responses highlight how the AI-generated feedback is perceived as beneficial for restructuring or enhancing one's work.

It may be insightful to examine the graph in relation to students' dimensions of feedback literacy, which, as Carless and Boud (2018) explain, include the ability to appreciate feedback, make informed judgments, and manage emotions to fully utilise all received information and stimuli to enhance learning. Students reported that the feedback they received was beneficial in supporting these dimensions (*The feedback was helpful for our self-esteem because it highlighted many positive aspects of our teaching unit; I must say that the positive feedback filled us with pride, reinforcing our belief that our work was solid and well-considered).*

Students emphasised the emotional and positive impact of feedback and its role in supporting *managing affect* (frequency = 4). They also highlighted *accuracy* (frequency = 7) and the function of promoting *reflection* (frequency = 5), which enabled them to engage more deeply with the feedback, learning to evaluate the strengths and weaknesses of their work. Additionally, they noted that the feedback provided specific *suggestions* and, as previously mentioned,

guidance for the *integration* of improvements into their own work. This process allowed students to experience the generative and transformative potential that timely feedback on their practice can foster (Laici, 2021).
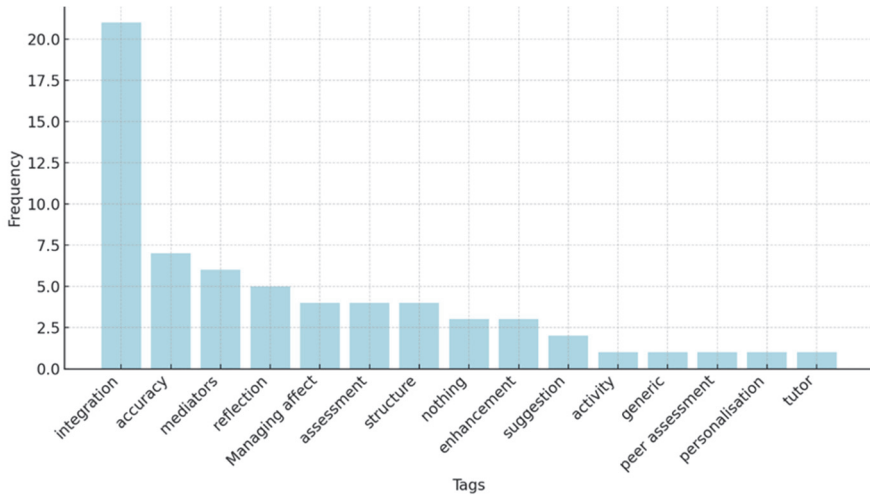

Fig. 7 - Strengths of feedback given by Chat GPT: assigned tags frequency

This interpretation is corroborated by responses to an earlier closed-ended question on a 5-point Likert scale, where students were asked to rate the usefulness of the feedback they received from ChatGPT. The majority rated it as either very useful or fairly useful, with only one student rating it as completely useless (see Fig. 8).
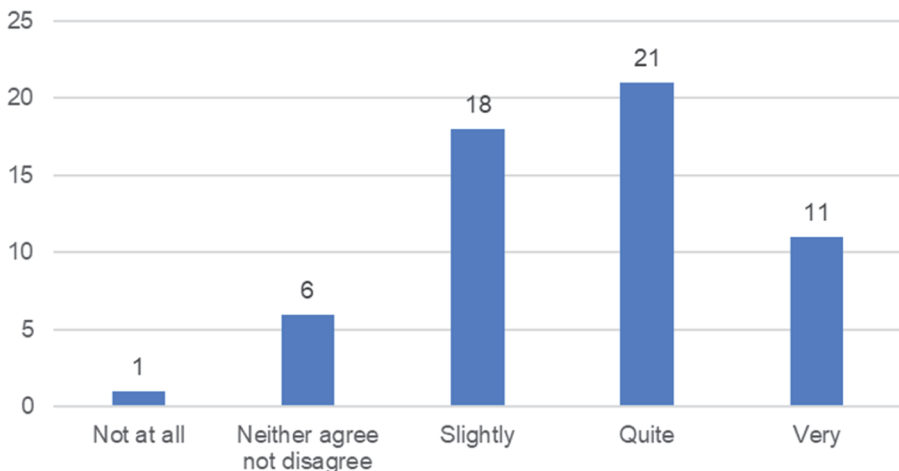

Fig. 8 - Feedback usefulness levels

An analysis of the co-occurrences between responses to the closed question and the tags assigned to open-ended responses on the strengths of the feedback (see Fig. 9) reveals an interesting pattern: most of those who selected the 'very' useful level attribute this usefulness specifically to the generative quality of the feedback, particularly in terms of *integration* with their task.

The tagging of responses regarding the weaknesses of ChatGPT's feedback highlights a perception of certain evident limitations in the comments provided (see Fig. 9).
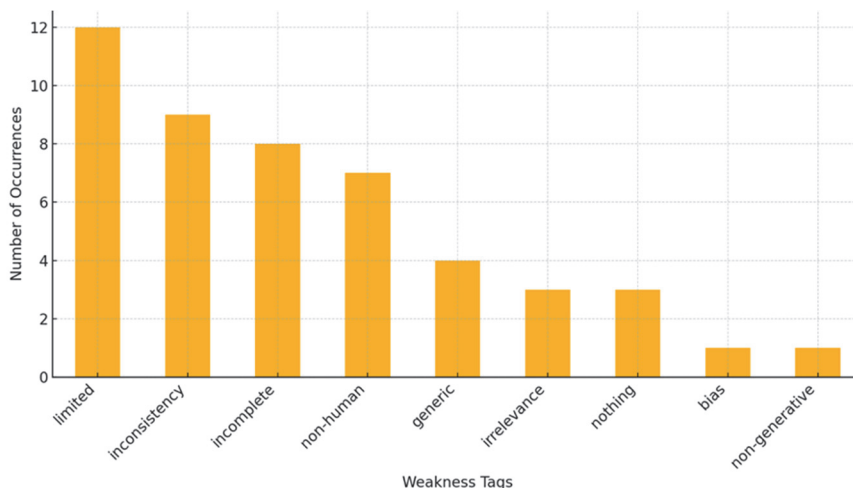

Fig. 9 - Weaknesses of feedback given by Chat GPT: assigned tags frequency

This is aligned with the comments associated with the "non-human" tag concerning the chatbot's design.

Students perceive the feedback provided by the tool as *incomplete* and, at times, *inconsistent*. Generally, respondents expressed a lack of appreciation for the excessive coldness and detachment in the AI-generated feedback, to the extent that they attributed the absence of certain insights or behaviors – characteristic of human feedback – to limitations of the tool. Indeed, students integrated ChatGPT into the educational ecosystem and expected a more contextualised type of feedback, one that could consider elements only implicitly present in their design but not fully defined, making them elusive for an AI system (Pentucci, Sarra and Laici, 2023). These observations are summarised under the *non-human* tag.

A closer examination of responses reveals that students found the feedback too impersonal, lacking the human connection and interaction they would typically receive from a teacher. The comments were often described as rigid

or overly *generic*, unable to adapt to the specific pedagogical needs of the context.

It is noteworthy how students reflect on the process, viewing ChatGPT as a human-like agent and expecting responses or reactions that are, in reality, unachievable for an artificial intelligence. This aligns with a perspective currently explored in the literature, which examines how the conversational and dialogical capacities of LLMs can foster a sense of human interaction, generating related expectations and reactions (Reiss, 2021). Students, for instance, commented: "*A weak point might be the lack of subjective value; it does not consider the individual members of the class, particularly the specific needs of individual pupils, which a teacher would certainly take into account*".

Analysing the co-occurrences between the tags assigned to strengths and weaknesses (see Fig. 10), we observe that students who appreciated the proactive potential of the feedback – its ability to encourage revisions and additions – also expressed concerns about its *inconsistency*.

Fig. 10 - Heatmap of co-occurence between tags

While the tool can suggest or prompt improvements, it cannot directly integrate these changes. Students recognise the feedback as accurate, yet are aware of its limitations, as it addresses only certain broad or partial aspects.

It appears, therefore, that using ChatGPT in peer feedback processes is indeed beneficial, particularly as a reinforcement tool that activates students' self-reflection on their own work (Lee et al., 2023).

Summarizing, decisions on what to modify or how to enhance the task ultimately rest with the student, who recognises that they have access to a precise analysis of their work, yet one that lacks human insight, contextual sensitivity, and an understanding of the nuanced subtext inherent in complex educational settings (Baidoo-Anu and Ansah, 2023).

4.3 ChatGPT and Professional Development

Students were asked if they foresaw using ChatGPT as a tool to improve their professional development in the near future and 77% would use it. The absolute frequencies of the tags related to the open-ended question "*For which activities would you use it*?", as evidenced by fig. 11, reveal a multifaceted reality: the analysis of the two main tags shows that the majority of future educational professionals find ChatGPT particularly useful for *design* (N=25), followed by *reviewing educational activities* (N=9).
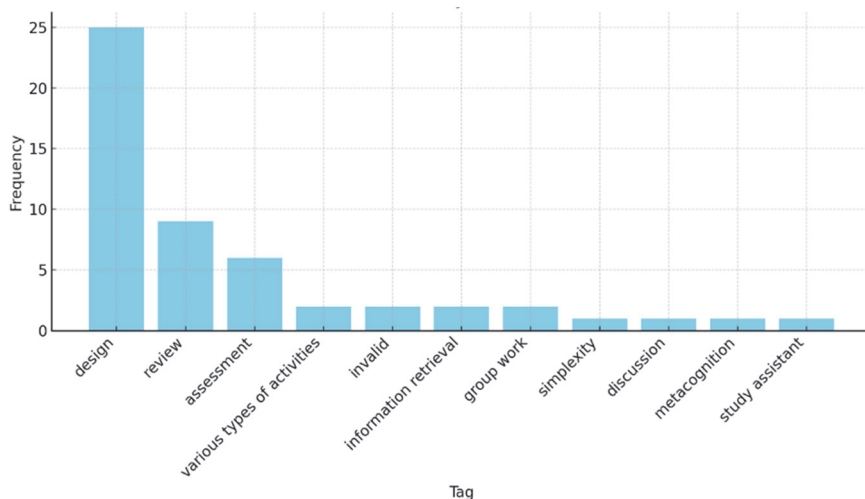

Fig. 11 - *Use of ChatGTP by teachers and educators: assigned tags frequency*

First and foremost, future teachers and educators see ChatGPT as a tool that can facilitate the implementation of *design*. It is useful for creating educational

pathways and supporting the learning design of educational actions, although they are aware that it requires constant supervision and process control, recurring element in previous student reflections (Rossi and Pentucci, 2021). In addition, ChatGPT is often seen as a basis from which to integrate activities based on specific educational contexts, aspect that complements the design view that regarded this as an aspect of weakness. (*It could be useful to provide starting points for activities, but it definitely requires reflection before bringing such activities into the classroom*; *I wouldn't use it to design something from Scratch, but rather to get inspiration or advice on activities proposed by the teacher*).

Secondly, ChatGPT is seen a useful tool for *reviewing* educational designs, materials, and interventions created by educational professionals. It serves as a means of checking, correcting and improving the different resources and pathways prepared by teachers and educators before they are delivered to students (*I would use it for comparison with what I, as an educator, have already designed for my class as a way to compare my ideas*; *To refine the details of a design*; *As a comparison element to achieve an even more complete design*). Furthermore, the tool could support students in reflecting critically on metacognitive learning processes, promoting a more conscious and reflective approach to studying (*It could play the role of an 'external opinion' to help students review their work*).

From the overall reading of the interventions, it emerges the that ChatGPT can function not only as a simple support tool, but also as a potential educational partner capable of integrating smoothly and dynamically into different stages of the educational experience, provided that the educational professional always maintains control over the technology: "*This must always be done under the critical and vigilant eye of the teacher, who must be able to recognise when ChatGPT's suggestions make sense and when they are incompatible with the objectives or needs of the class*".

ChatGPT is seen as a tool that increases the efficiency of teachers' and educators' work, providing solid support for design, thus optimising time and improving the quality of the educational materials produced.

The last two questions analysed asked the students whether, in their future professional role as educators and teachers, they intended to have their students use ChatGPT and 61% would propose it to students. Although this figure is lower, it is not too far from the hypothesis regarding the use of the chatbot by teachers and educators. This suggests that generative AI is perceived by the young generation in training as a tool that can be fluidly and dynamically integrated with others in the educational ecosystem of the classroom.
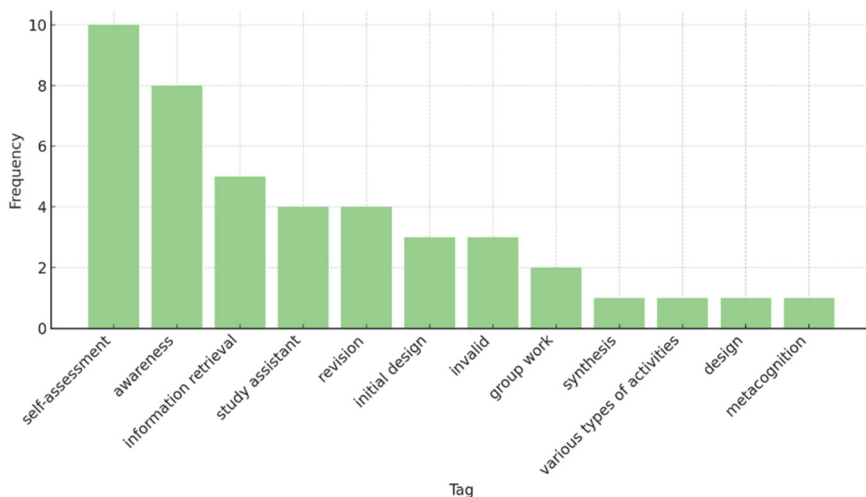
Fig. 12 - Use of ChatGTP by students: assigned tags frequency

Analysing the absolute tag frequencies of the four main tags in response to the open-ended question "*What activities would you have students use it for?*", as evidenced by fig. 12, the majority of future educators and teachers see ChatGPT as a *self-assessment* tool for their students (N = 10), followed by a means of raising *awareness* of the tool itself (N = 8).

Firstly, future educational professionals see the chatbot as a self-assessment tool that can help students to reflect metacognitively on their learning, develop a greater awareness of their gaps and progress, and thus integrate and improve their work, while strengthening their autonomy and self-evaluation skills (*To evaluate the work they have done and for self-assessment*; *As a means to compare the work*; *To correct their assignments or get some suggestions*; *As an 'evaluation' to allow them to catch all the aspects they might have missed*). This is achieved through specific activities and ways of using the tool proposed by the teacher (*Activities of the type 'find the error' or 'make a critical comment on this completed exercise'*).

In addition, if used as a meta-reflective tool, the chatbot would offer the possibility of critical reflection and awareness of the tool itself, its potential and its limitations (*I would use it as a tool to allow students to critically analyse what they find online, even in the field of mathematics. I would like to convey the message that the chatbot is not an oracle, but a good tool to learn from mistakes*; *Solve exercises, ask questions on topics, but then ask the students where the mistakes are*).
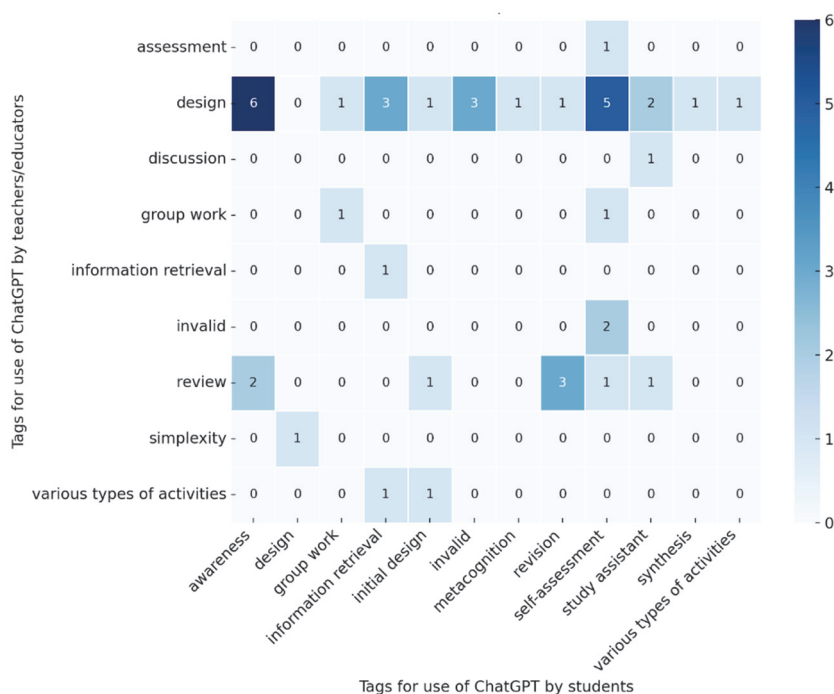
Fig. 13 - Heatmap of co-occurence between tags

If we analyse the heatmap of co-occurrences between the answers "*Would you use ChatGPT in an educational context as a tool for teachers/educators?*" and "*Would you let your students/learners use ChatGPT?*" (fig. 13), we see a strong co-occurrence (N = 6), indicating ChatGPT as a key design tool in the hands of educational professionals and as a means of raising students' awareness of the tool itself, with the aim of using it critically to support their learning. Another significant co-occurrence (N = 5) is between ChatGPT as a design tool in the hands of teachers and as a valid self-assessment tool for students. This relationship seems to reflect an educational model centred on self-regulation students are asked to measure their progress independently through the chatbot. In this way, ChatGPT would become a facilitator of reflective learning, fostering autonomy and awareness of their educational pathway (van Velzen, 2015).

The awareness that it has become a tool of everyday use emerges; ChatGPT consequently requires careful attention from the educational world in terms of developing digital literacy in the younger generations, ensuring its critical,

responsible and ethical use (*Moreover, I believe that it is not wrong for students to approach AI because it is part of the world we live in and it will only expand*).

5. Conclusions

The use of ChatGPT in educational contexts and in teacher training seems to confirm what the international literature points out about AI-based tools: they can be proactive partners and agents within educational ecosystems.

In our view, the primary interest emerging from this study centres on students' perceptions of the tool within the human/non-human dichotomy and its transformative potential, particularly in relation to students' reflections across the different phases of the work.

The transformative process starts when the tool is used for design and feedback. At this stage, students expect human-like reactions or responses from the chatbot. Instead, they encounter a non-humanity that initially emerges as a weakness: in performing the didactic actions, the students notice criticalities such as coldness, lack of contextualisation, generalisations and standardisations. At a later stage of the research activity, when the students were asked to reflect on their future, they distanced themselves from the highlighted strengths and weaknesses. This generated an initial realisation: the tool is useful and potentially applicable in educational contexts, but human intervention remains indispensable to fill the gaps outlined above.

This encourages us to continue a research approach where AI can act as a support agent for educational and didactic practices, but it needs to be integrated with pedagogical awareness and sensitivity, strengthening an AI culture capable of balancing efficiency and humanity (Elliot, 2021; Panciroli and Rivoltella, 2023).

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.

Bolter J.D. (2019*). The Digital Plenitude. The Decline of the Elite Culture and the Rise of New Media*. Cambridge, MA: TheMIT Press.

Bonaiuti G., Calvani A., Menichetti L., and Vivanet G. (2017). *Le tecnologie educative. Criteri per una scelta basata su evidenze.* Roma: Carocci.

Braun V., Clarke V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2): 77-101. DOI: 10.1191/1478088706qp063oa.

Braun V., Clarke V. (2012). Thematic analysis. In H. Cooper P. M. Camic D. L. Long A. T. Panter D. Rindskopf and K. J. Sher (Eds.). *APA Handbook of Research Methods in Psychology: Vol. 2. Research Designs* (pp. 57-71). Washington, DC: American Psychological Association.

Braun V., Clarke V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4): 589-597. DOI: 10.1080/2159676X.2019.1628806.

Capolla L.M., Giannandrea L., Gratani F., Pentucci M. and Rossi P.G. (2024). Rethinking and formalizing initial teacher training on learning design for and in uncertainty. *Frontiers in Education*, 9, 1268936. DOI: 10.3389/feduc.2024.1268936.

Carless D., Boud D. (2018). The development of student feedback literacy enabling uptake of feedback. *Assessment and Evaluation in Higher Education, 43*(8): 1315-1325.

Coeckelbergh M. (2020). *AI Ethics*. Cambridge MA: The MIT Press.

Deng Y., Liao L., Zheng Z., Yang G.H., Chua T. (2024). Towards human-centered proactive conversational agents. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022* (pp. 807-818). DOI: 10.1145/3626772.3657843.

Elliott A. (2021). *La cultura dell'intelligenza artificiale*. Torino: Codice Edizioni.

Følstad A., Araujo T., Law E.L.C., Brandtzaeg P.B., Papadopoulos S., Reis L., Baez M., Laban G., McAllister P., Ischen C., Wald R., Catania F., Meyer Von Wolff R., Hobert S. and Luger E. (2021). Future directions for chatbot research: An interdisciplinary research agenda. *Computing*, 103(12): 2915-2942. DOI: 10.1007/s00607-021-01016-7.

Gratani F., Capolla L.M., Giannandrea L., Screpanti L. and Scaradozzi D. (2023). Facilitating feedback at university using AI-based techniques. In *HELMeTO 2023-Book of Abstracts* (pp. 157-159). Roma: STUDIUM.

Gupta P., Ding B., Guan C. and Ding D. (2024). Generative AI: a systematic review using topic modelling techniques. *Data and Information Management*, 8(2), 100066. DOI: 10.1016/j.dim.2024.100066.

Hwang G.J., Chang C.Y. (2021). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, 1-14. DOI: 10.1080/10494820.2021.1952615.

Krippendorff K. (2018). *Content analysis: An introduction to its methodology*. London: Sage.

Laici C. (2021). *Il feedback come pratica trasformativa nella didattica universitaria*. Milano: FrancoAngeli.

Lee H., Phatale S., Mansoor H., Lu K.R., Mesnard T., Ferret J., Bishop C., Hall E., Carbune V. and Rastogi A. (2023). RLAIF: Scaling reinforcement learning from human feedback with AI feedback. In *Proceedings of the 41st International Conference on Machine Learning* (PMLR 235). Vienna: PMLR.

Liao L, Yang G., Shah C. (2023). Proactive conversational agents in the post-ChatGPT world. *SIGIR,* 23: 23-27.

McNamara D.S., Arner T., Butterfuss R., Fang Y., Watanabe M., Newton N., McCarthy K.S., Allen L.K. and Roscoe R.D. (2023). iSTART: Adaptive comprehension strategy training and stealth literacy assessment. *International Journal of Human-Computer Interaction*, 39(11): 2239-2252.

Nicol D. (2020). The power of internal feedback: exploiting natural comparison processes. *Assessment & Evaluation in Higher Education*, 46(5): 756-778. DOI: 10.1080/02602938.2020.1823314.

Nicol D., S. McCallum (2022). Making Internal Feedback Explicit: Exploiting the Multiple Comparisons That Occur during Peer Review. *Assessment & Evaluation in Higher Education* 47(3): 424-443. DOI: 10.1080/02602938.2 021.1924620.

Nieminen J.H., Carless D. (2023). Feedback literacy: A critical review of an emerging concept. *Higher Education*, 85(6): 1381-1400. DOI: 10.1007/s10734-022-00895-9.

Packin N. G., Lev-Aretz Y. (2018). Learning algorithms and discrimination. In W. Barfield, U. Pagallo (eds.). *Handbook on the law of artificial intelligence* (pp. 88-113). Cheltenham MA: Edward Elgar Publishing.

Panciroli C., Rivoltella P.C. (2023). Pedagogia algoritimica. *Per una riflessione educativa sull'Intelligenza Artificiale*. Brescia: Scholè.

Pang T.Y., Kootsookos A., Cheng C. (2024). Artificial intelligence use in feedback: a qualitative analysis. *Journal of University Teaching and Learning Practice*, 21(6): 1-18. DOI: 10.53761/40wmcj98.

Pentucci M., Laici C. (2023). Peer feedback and peer review in higher education: mirroring and transformative practices in active learning processes. In *EDULEARN23 Proceedings* (pp. 790-799). Siviglia (ES): IATED.

Pentucci M., Rossi P.G., Capolla L. (2023). Designing regulation in action to manage the unforeseen in teaching-learning contexts. In *EDULEARN23 Proceedings* (pp. 1151-1158). IATED.

Pentucci M., Sarra A., Laici C. (2023). Feedback to align teacher and student in a digital learning ecosystem. *Education Sciences & Society, 14*(1): 242-260. Milano: FrancoAngeli.

Pentucci M. (2018). *I formati pedagogici nelle pratiche degli insegnanti*. Milano: FrancoAngeli.

Perera P., Lankathilaka M. (2023). AI in higher education: A literature review of chatgpt and guidelines for responsible implementation. *International Journal of Research and Innovation in Social Science*, *7*(6); 306-314. DOI: 10.1080/10494820.2021.1952615.

Prananta A.W., Megahati S., Susanto N. and Raule J.H. (2023). Transforming education and learning through ChatGPT: A systematic literature review. *Jurnal Penelitian Pendidikan IPA*, 9(11): 1031-1037. DOI: 10.29303/jppipa.v9i11.5468.

Reiss M.J. (2021). The Use of AI in Education: Practicalities and Ethical Considerations. *London Review of Education*, *19*(1), 1.

Rivoltella P.C., Rossi P.G. (2019). *Il corpo e la macchina. Tecnologia, cultura, educazione*. Brescia: Scholé.

Rossi P.G., Pentucci M. (2021). *Progettazione come azione simulata. Didattica dei processi e degli eco-sistemi*. Milano: FrancoAngeli.

Rossi P.G., Pentucci M., Fedeli L., Giannandrea L. and Pennazio V. (2018). Dal feedback informativo al feedback generativo. *Education Sciences and Society*, *9*(2): 83-107.

Tam A.C.F. (2024). Interacting with ChatGPT for internal feedback and factors affecting feedback quality. *Assessment & Evaluation in Higher Education*, 1-17. DOI: 10.1080/02602938.2024.2374485.

Sansone N., Bortolotti I., Fabbri M. (2021). Il peer-assessment nella formazione insegnanti: accorgimenti e ricadute. *Education Sciences & Society*, 2: 444-460.

Sansone N., Fabbri M., Bortolotti I. (2023). Reflective practices in teacher education for and through technologies. *Formazione & insegnamento*, 21(1): 239-246. DOI: 10.7346/-fei-XXI-01-23_29.

van Velzen J.H. (2015). Are students intentionally using self-reflection to improve how they learn? Conceptualising self-induced self-reflective thinking. *Reflective Practice*, 16(4): 522-533. DOI: 10.1080/14623943.2015.1064378.

Walter Y. (2024). Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, 21(1), 15. DOI: 10.1186/s41239-024-00448-3.

Wan T., Chen Z. (2024). Exploring generative AI assisted feedback writing for students' written responses to a physics conceptual question with prompt engineering and few-shot learning. *Physical Review Physics Education Research*, *20*(1), 010152.

Williamson B., Macgilchrist F., Potter J. (2023). Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, 48(1): 1-5.

Winstone N., Carless D. (2019). *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach* (1st ed.). London: Routledge. DOI: 10.4324/9781351115940.