MIND IN MIND LAB: EVALUATION TOOLS AND CRITICAL ANALYSIS

MIND IN MIND LAB: STRUMENTI DOCIMOLOGICI E ANALISI CRITICA

Martina Albanese
University of Palermo
martina.albanese@unipa.it

Elisabetta Fiorello
University of Palermo
elisabetta.fiorello@unipa.it

Giuseppa Compagno
University of Palermo
giuseppa.compagno@unipa.it

Double Blind Peer Review

Citazione
Albanese, M., Fiorello, E., & Compagno, G. (2024). Mind in mind lab: evaluation tools and critical analysis *Giornale Italiano di Educazione alla Salute, Sport e Didattica Inclusiva*, 8(2), Edizioni Universitarie Romane

Doi:
https://doi.org/10.32043/gsd.v8i3.1135

gsdjournal.it
ISSN: 2532-3296
ISBN 978-88-7730-494-0

ABSTRACT
This paper intends to outline possible paths to guide university students to a critical use of ChatGPT and to explore the software's limits and potential as an educational tool to be integrated in learning activities and in the processes of searching for reliable sources, through metacognitive reflection. For process and product assessment, the use of the Meta-reflection Form on the use of ChatGPT plays a fundamental role.

Il presente contributo intende tracciare percorsi possibili che guidino gli studenti universitari ad un uso critico di ChatGPT e che permettano di esplorare limiti e potenzialità del software come strumento educativo da integrare nelle attività di apprendimento e nei processi di ricerca delle fonti attendibili, attraverso una riflessione metacognitiva. Per la valutazione di processo e di prodotto assume un ruolo fondamentale l'utilizzo della Scheda di Meta-riflessione sull'utilizzo di ChatGPT.

KEYWORDS
Laboratorio; chatGPT; intelligenza artificiale; valutazione.
Laboratory; chatGPT; artificial intelligence; assessment.

Received 30/04/2024
Accepted 19/06/2024
Published 24/06/2024

Introduction[1]

In recent years, the field of Educational Sciences has been strongly influenced by emerging theoretical and experimental insights from neuroscience, alongside the development of technologies based on artificial intelligence (AI) systems. The fusion of AI studies with research on human cognition has led to a new debate in contemporary scientific discourse concerning the potential impact of such technologies on shaping the cognitive learning processes of future citizens (Drivas & Doukakis, 2022).

While it's acknowledged, as Minello (2011) points out, that the cognitive sciences have yet to produce a comprehensive paradigm applicable to individual educational situations, in a scenario where advancements in new technologies permeate numerous fields of knowledge, the epistemological framework of Neuroscience can serve as a useful lens for interpreting educational phenomena in their intricate, often nonlinear relationship with artificial intelligence.

Starting from the premise that Cognitive Neuroscience refers to an interdisciplinary research field aimed at understanding the brain mechanisms involved in various cognitive domains (e.g., perception, language, reasoning, emotions, executive functions, motor skills), and hence, learning (Berthier et al., 2018; Damiani, Santaniello, Gomez Paloma, 2015), it could be inferred that by leveraging some of its derived principles, it is possible to control and guide the variables at play in the interplay with artificial intelligence.

Simultaneously positing the uniqueness of individuals and educational contexts, it is believed possible to draw connections between brain function on one hand and the potentials of artificial intelligence on the other, starting from the idea that humans should be recognized as a "bio-psychosocial-spiritual unit, in which the inseparability between body and thought confers upon them a character of uniqueness" (Vinci, 2019).

Based on these premises, it is desirable for educators, equipped with critical minds and field experiences, to explore new paths through practices supported by updated scientific knowledge; this would allow them to expand their professional skills with new abilities and integrate digital technologies into teaching-learning processes from a neuroeducational perspective. In line with this vision, AI can

[1] This paper is a co-authored work of the three authors; however, Martina Albanese is the author of the introduction and conclusion, Elisabetta Fiorello is the author of par. 2, 2.1 and 3.2, Giuseppa Compagno is the author of par. 1, 3.1.

become a means to facilitate knowledge transmission, creativity acquisition, memorization, and cognitive flexibility, enriching and complementing educational actions without limiting them (Minello, 2020).

It should be emphasized that learning, as so conceived, occurs in various modalities and using different tools, methodologies, and technologies, always encapsulating a cognitive component related to the acquisition and consolidation of information and concepts in memory systems, associated with well-defined emotional-sensory stimuli; thus, it transcends the specificity of the means employed.

This contribution aims to apply the aforementioned discussion to the specific theme of using ChatGPT in educational contexts, followed by presenting a laboratory experience conducted in a university setting where Chat GPT was utilized for constructing a scientific poster and a Meta-Reflection Scheme on the critical use of the software.

1. L'uso di ChatGPT nei contesti educativi

Over the past decade, the widespread use of AI and OpenAI software has permeated many sectors, including education. Particularly ChatGPT, a conversational chatbot recently developed by OpenAI, has, since its launch on November 30, 2022, become the fastest-growing user application in history, generating as much excitement as discontent within school and academic settings (Lo, 2023).

The features of the chatbot, developed according to a generative artificial intelligence model[2], have garnered the attention of the education sector as they essentially revolve around the possibility of improving learning processes by serving ChatGPT's ability to provide consistent, systematic, and informative responses.

The benefits that the use of artificial intelligence can bring to educational contexts have been emphasized especially by that orientation of the scientific literature that, by virtue of the innovative scope of this technology, has sometimes taken positions of uncritical fervor (Chen & Lin, 2020). The advantages that the authors have repeatedly identified refer, among others, to the possibility of students' personalized learning experiences (Oranga, 2023). Indeed, ChatGPT, by adapting

[2] Generative artificial intelligence (generative AI) is a type of artificial intelligence that can create new content and ideas, including conversations, stories, images and music, by simulating the use and processing of human language.

content and explanations to the needs and pace of its interlocutor, modulates its outputs and responses according to the level of understanding of the subject interacting with it.

From a pedagogical perspective, the effect of the use of chatbots on learning processes would result in the real opportunity to respond to students' individual needs, provide them with immediate feedback, and facilitate understanding of complex concepts, promoting active participation and cognitive advancement (Zhai,2022). Furthermore, by adopting this approach, students can benefit from ChatGPT's assistance throughout the day, even beyond school or university hours. This means they have a learning tool readily available to provide immediate and accessible answers for assignments, quizzes, and questions, aiding them in understanding mistakes and rectifying them promptly (Buchberger, 2023).

Another set of advantages associated with integrating AI into educational contexts stems from the vast wealth of information that applications like ChatGPT offer, spanning various topics and disciplines. Learners can tap into the extensive data and knowledge stored in the databases these software access, receiving assistance across a wide array of scientific fields and domains.

This potential horizon is mirrored, in turn, in the much-debated opportunity to utilize ChatGPT for foreign language acquisition and enhancing general language proficiency, including in one's native language. This includes grammar corrections, suggesting synonyms, and explaining idiomatic expressions.

Especially in upper secondary and university education, AI would emerge as a valuable tool to foster the advancement of students' writing abilities. This would impact their written communication skills and effectiveness in crafting scientific texts and materials, particularly within university settings (Montenegro-Rueda et al., 2023).

Moreover, literature has highlighted how ChatGPT can facilitate not only individual learning but also collaborative learning, aiding users in working together on projects or tasks involving problem-solving. It also fosters the creation of a cooperative and inclusive atmosphere.

In group settings, particularly, utilizing a chatbot proves effective in reducing anxiety levels among introverted students or those less inclined to interact actively. This stimulation of creating spaces to ask questions and seek assistance from both peers and teachers is instrumental in promoting engagement and participation (Wang et al., 2023).

As Oranga (2023) noted, ultimately, the use of AI software would be cost-effective in the sense that it would eliminate the need to employ additional costs to hire tutors and human support professionals, benefiting those who use it economically.

This position, however, should be considered cautiously because of the strong human component inherent in the tutoring principle that can neither be ignored nor replaced. Exactly here lies the meaning of the educational dimension that is person-centered and cannot be supinely bypassed by techno-digital substitutions that do not consider aspects such as pedagogical intentionality and participation in the educational relationship.

Even though, based on the reported evidence, ChatGPT shows enormous potential in improving the efficiency of learning processes and for providing personalized educational support, both to students and teachers, it is essential to consider and highlight-with equal zeal-the risks and limitations associated with the use of these technologies, from a pedagogical perspective.

Most of the concerns that have been raised, in the research field, about possible disadvantages and limitations that Open AI software presents, relate to the fear that these technologies may progressively erode students' critical thinking (Wang et al., 2023; Fuchs et al., 2023). The possible reduction in the judgment skills of ChatGPT users would be ascribable to a structural defect of the chatbot, which is known not to possess true comprehension and analytical reasoning skills, but rather generates responses that are not always accurate or well-reasoned, based on patterns in the data with which it has been programmed and on which it can draw. This type of obviously limited and partial knowledge can also produce contextually inappropriate or meaningless outputs due to a lack of deep understanding of the broader context outside the conversation by the chatbot.

Furthermore, while ChatGPT can offer information, it lacks the ability to verify its accuracy, necessitating users to critically assess the information provided periodically. These limitations result in slowdowns and difficulties in time management among students, as well as confusion and misinformation (Halaweh, 2023).

Despite AI software developers having access to vast amounts of information, their knowledge may not always be current and often lacks expertise in specialized or niche topics. Additionally, the implementation of ChatGPT in educational settings and beyond raises several ethical and regulatory concerns that have yet to be adequately addressed and resolved. In a broader context, the recording and archiving of conversations with chatbots raise privacy issues and concerns

regarding the sharing of sensitive information. Moreover, the repeated recognition of the risk of ChatGPT being misused for generating unethical content, such as fake news, spam, or scams, is noteworthy. In university settings specifically, the use of such software has been linked to high risks of plagiarism in writing papers, theses, and other written assignments (Javaid et al., 2023; Strzelecki, 2023).

Considering all previous considerations, it is acknowledged that chatbots can serve as a supportive tool capable of enhancing students' learning experiences, albeit cautiously. However, the most effective and least harmful implementation of these technologies occurs when they are utilized under teacher supervision. The proposed solution to mitigate both blind enthusiasm and irrational distrust towards the use of AI in educational settings is to introduce targeted educational pathways. These pathways encourage students to use ChatGPT consciously and responsibly, providing them with a comprehensive understanding of its limitations and potential.

While longitudinal data on the impact of such technology on knowledge and learning processes are lacking, it is crucial to begin incorporating targeted training actions into educational programs immediately. These actions aim to prevent the potential drifts associated with the abuse or misuse of artificial intelligence software.

2. The *Mind in Mind* Lab

During the second semester of the present academic year 2023-2024, the Mind in Mind Lab involving 160 students of the Evaluation teaching & lab of the Degree Course in Pedagogical Sciences of the University of Palermo is being carried out. The laboratory, lasting 30 hours, involves the realization of a path of work in a laboratory set-up, distinguished in phases that leads to the realization of an educational project proposal accompanied by effective assessment tools with respect to the development of some neuro-oriented skills.

The initial phase of the laboratory, central to this contribution, focuses on the critical utilization of ChatGPT to explore its limitations and potential as an educational tool integrated into learning activities and processes of sourcing reliable information. This entails activating a metacognitive reflection process. Therefore, a key objective is to guide students towards critically engaging with artificial intelligence, debunking its potentialities, highlighting implicit risks, and providing assessment tools to navigate the activated metacognitive journey. As

outlined in the preceding section, the sudden emergence of ChatGPT in 2022 caught the educational community off guard and elicited numerous concerns (Huang et al., 2023). Immediate criticism from colleges and universities centered on the fear that this tool would be prone to indiscriminate plagiarism and could generate texts undetectable by software, potentially leading to students illegitimately appropriating written work (Susnjak, 2023).

For this reason, the *Mind in Mind lab* that is the subject of this paper was constructed by assuming a methodological and operational posture apt to significantly improve the critical use that can be made of AI-based chatbots, especially in university educational contexts; without, however, forgetting the risks and problems inherent in their implementation, which assume in this perspective the role of theoretical core around which the experimental action revolves.

2.1 The methodology and work steps

Tapping into the knowledge that "when we feel socially supported, our executive brain functions improve" (Guillén, 2021), the proposed workshop activities were set up by taking advantage of the peculiarities of Student Team Learning (STL) methodology (Slavin, 1991).

The realization resulting from the use of this approach to the learning process is that each student knows that a group of peers supports his or her efforts; in fact, the success of the group requires that all members do their best (Slavin, 1991). The main feature of this methodology involves dividing students into heterogeneous groups. Another distinguishing component turns out to be the increase in student motivation, which, according to the U.S. scholar, would increase individual sense of responsibility as well as the performance of each group member (Benzato & Minello, 2002). The context of positive competition, which is established, entails the fact that all students would be pushed to significantly improve their performance.

Among the various STL techniques developed, we chose to use the Jigsaw technique that takes advantage of task specialization. Jigsaw involves the creation of heterogeneous groups of 3-5 members (Jigsaw groups); each student is assigned a part of the task to make him or her an "expert" on that segment by becoming responsible both for teaching that information to the other group members and for furthering the information received from the other group members on the other "sections" of the lesson (Mattingly & VanSickle, 1991).

Following this approach, the 160 attending students were divided into 32 heterogeneous groups. In the start-up phase of the activity, the teacher made explicit the theoretical assumptions that form the background of the activity and shared the guidelines for setting up the work, achieving the specific objective of the activity to be developed, evaluation criteria and evaluation tools.

The specific objective of the activity involved the creation of a scientific poster identifying the theoretical constructs and scientific evidence related to the acquisition and development of certain neuroscience-based skills (Albanese & Compagno, 2022; 2023), using AI to search (and refute) sources.

Each group was assigned a competency that looks at one of the following neuroteaching areas: Transversal (general brain functioning, general methodological-didactic strategies, Soft skill development, Learning environment/setting), Socio-Emotional-Affective, Cognitive, Linguistic-Communicative, and Praxical-Motor. Starting from the consultation of OpenAI's chatbot, chatGPT, students cooperate to collect sources and create the poster according to the cooperative action described in the table below:

Goals and Phases of Cooperative Work			
Prerequisites: knowledge and study of neuro-educational areas; knowledge of the basic working mechanism of ChatGPT; knowing how to use digital tools.			
General objective: cooperation and distribution of individual responsibility with respect to the task; critical and reasoned use of ChatGPT; knowing how to refute sources.			
Specific Objective: Creation of a scientific poster on the theoretical framework supporting the assigned neuro-educational competence.			
	What the teacher does	*What the student does*	*Note*
Step 1 - 15 min. Before the activity	Exposes the theoretical framework and gives the assignment; allocates within the groups the specific tasks (segments); accommodates doubts and questions.	Listens to directions, forms the group, and chooses the leader for discussion with the teacher	Jigsaw Segments: Using chatbot, setting up poster graphics, refuting and consulting other sources, developing poster content, leading for teacher-group communication
Phase 2 - 30 min. activity	Monitor the first phase	Search for poster creation information on chatGPT	
Phase 3 - 30 min. activity	Provides and exposes the Meta-reflection Outline on critical use of GPT chat; accommodates concerns and questions	Adopts and fills out the tool provided by the teacher to initiate metacognitive reflection and evaluation of chatbot use.	

Step 4 - 15 min. activity	Monitor the second phase	Expert Groups meet to exchange information	Each expert compares with other experts in the field to exchange strategies and ways of working
Step 5 - 30 min. activity	Monitors and supports the development of the third phase	"Mother groups" rebuild and produce science poster	
Step 6 - 60 min. After the activity	Evaluates products and adjusts peer review	Classroom display of the created poster; self-assessment.	Each group evaluates the products created according to evaluation criteria provided by the teacher

Table 1 (Goals and stages of cooperative work)

Erfiani and Neno (2018[3]) highlight how Jigsaw helps improve mutual trust among students, gives space for the emotional and relational component, and ensures, from an inclusive perspective, equal participation of all students through peer tutoring (scaffolding).

The criteria for evaluating the poster concern: the reliability of the sources, the relevance of the theoretical background to the competence assigned to the group, the identification of the method(s) found in the contributions analyzed, the presence of all the characteristic elements of the poster (introduction, background, methods, analysis, and conclusion).

In the evaluative perspective, for the purpose of the process and product evaluation reached by the students with respect to the development of the first phase of the workshop course, the use of the Meta-reflection Form on the use of ChatGPT, which was specially constructed and whose identified levels of analysis concern (Table 2), assumes a fundamental role: the ease or difficulty in formulating the prompts, (whether and how many times it was necessary to enter different prompts to obtain the desired response); the effectiveness of the software in terms of time managing; the veracity and accuracy of the information obtained; and the relevance to the prompt of the responses developed by the software.

[3] Erfiani, Y.P.F., Neno, H. (2018), The Effect of Jigsaw Method to Improve EFL Students' Vocabulary Ability, *Metathesis: Journal of English Language Literature and Teaching,* Vol. 2, No. 2, pp. 171-183.

Analysis criteria	Guided stimulus questions	Hint	Examples
Prompt formulation (1)	Have I made the context explicit? Did I provide one stimulus at a time? Did I provide examples? Did I exceed 4,000 words? Did I use the affirmative form?	Explain the context/target of reference. Chained prompting: break down complex requests into intermediate steps, different but related prompts. One-shot prompting: provide along with a request, some sample content. Do not exceed 4096 tokens (about 4000 words in the input), excess text is ignored. Use clear verbs and affirmative form. Explicit the language register (high/academic/children's comprehension) that Chat GPT should adopt	Write a fairy tale for 4-year-olds. Write a fairy tale for 4-year-olds - insert the figure of a magic horse to the previous answer. Write a fairy tale for 4-year-olds following Munari's approach.
Verification of truthfulness of information (2)	Have I defined the prompt with the right information? Have I checked if the sources provided are existing?	Request exact sources. Check Google to see if the sources or information exists	Formulating an introduction on mirror neurons using scientific articles by Rizzolatti and colleagues Please provide sources published from 2019 to 2021
Checking the relevance of the output to the prompt (3)	Is there consistency between the answer given and the prompt used? Was I clear and accurate in the wording of the prompt?	Specify the information I want to obtain by better formulating the prompt. Avoid ambiguous language	Pubmed search how to make a diagnosis of PCI. Create a story for 4-year-olds starring animals of the savannah
Outcome/output provided (4)	Am I satisfied with the response received? Have I achieved the goal?	Refine information using other reliable channels. Improve the ability to provide accurate prompts	Use other reliable channels to search for sources (Google Scholar, Google Books, scientific journal and newspaper sites)
Difficulty employed (5)	Did I take a long time to formulate the prompt? Did I achieve the desired goal? Did I find it easy to use Chat GPT? Did I save	Practice improves with experience. Knowing the information search mechanism of Chat GPT	

	time by using ChatGPT to accomplish the task?		
Strengths on the use of AI			
Weaknesses on the use of AI			

Table 2 (Meta-reflection sheet on the use of ChatGPT)

3. Analysis of results

The evaluation tool used, the Meta-Reflection Form, combines and collects qualitative-quantitative data, and the data analysis is based on the use of thematic analysis techniques in which, student responses are first categorized and then analyzed with the aim of identifying patterns, trends, and recurring challenges that students encounter in using ChatGPT as an educational tool.

The research setting also allows students to experience a journey of building and validating an assessment tool. The latter, despite being used as a promoter of critical reflection with respect to the use of technologies that make use of artificial intelligence, is at the same time a tool for collecting data for the purpose of improving the evaluative-educational proposal.

In the initial phase of result analysis, we opted for a quantitative approach, utilizing descriptive statistical tools to examine both the sample characteristics and the data collected from the meta-reflection form. Regarding the sample characteristics, which comprised 160 students, the majority fell within the age group of 22 to 25 years (78.7%), with the remaining belonging to the age group of 25 to 36 years (21.35%).

In conjunction with biographical data, information on students' familiarity with ChatGPT usage was gathered for exploratory purposes. Of the participants, 21.4% reported prior usage of the chatbot, while 78.6% indicated no prior experience. Notably, among those familiar with the AI software, 50% stated they had utilized ChatGPT for academic learning purposes, while the other half had not. These findings, while based on a non-extensive sample allowing limited generalization, suggest that despite ChatGPT's innovative scope and widespread integration in educational settings, a significant portion of college students are yet to acclimate to, and consequently receive training in, utilizing the software.

In a second step, data from the analysis of the meta-reflection form were analyzed in percentage terms. The indicators in the form based on which the students conducted the meta-reflection process -- concerning the formulation of the prompt, verification of the truthfulness of the information, verification of the relevance of the output to the prompt, the outcome/output provided, and the difficulty employed, respectively -- are rated by the students on a three-class agreement scale. Specifically, the index of agreement could be expressed on a scale including the levels" yes," "somewhat," and "no."

Regarding the first criterion, related to the wording of the prompt, 57 percent of the students said they agreed that they had made explicit the context of the prompt provided, while 17.8 percent reported that they had not, as well as the remaining 25 percent were in the middle. Within the same criterion, 85.7% of the sample agreed that they had included one stimulus at a time, as reported in the prompts on the meta-reflection sheet. In contrast, with no disagreement position, 14.3% disagree with the question. In addition, 32.1% of students agree and thus say they included examples with which to accompany their prompts, when 46.4% did not and thus disagree. A high percentage of students (78.5%), in specifying whether they had used more than 4,000 words in formulating the prompt, agreed with the question. Similarly, most subjects (71.4%) assert that they had used the affirmative form in asking the chatbot questions and consequently agree.

Regarding the second criterion, pertaining to verifying the veracity of the information received as output, 78% believe they defined the prompt with the right information, 14.2% did so only partially, and 7.1% disagree.

Still, significantly, 85.7% say they checked whether the sources provided existed, 10.7% emphasize that they did so partially, while only 3.5% failed to do so.

Within the third criterion, having to do with checking the relevance of the output to the prompt provided, 89.2% say they agreed on the consistency between the answer provided and the prompt used, when only 10.7% did not.

From a metacognitive and meta-evaluative point of view, significant appears to be the data reporting that 82.1% of students believe they were clear and precise in formulating the prompt, while 17.8% state the opposite.

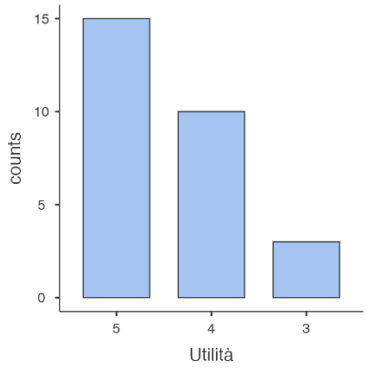
The fourth criterion, which investigated the dimension of the outcome/output provided by the software, central to the survey, revealed that 67.8% agreed with respect to being satisfied with the answer received and 31.1% disagreed instead.

Lastly, the results regarding the fifth criterion, which focuses on the degree of difficulty employed, describes a situation in which most students (53.5%) disagree that they spent a lot of time formulating the prompt, 35.7% partly and 10.7% assert that they spent a lot of time. At the same time, 75% say they achieved the desired goal and 85% express agreement by stating that they found it easy to use ChatGPT. Specularly, only 25% do not judge that they have achieved their desired goal and 14.2% express disagreement, having not found the use of ChatGPT easy.

The last phase of quantitative investigation involved analyzing the degree of agreement/disagreement with respect to statements about the use of Chat GPT for educational purposes following the experience in which the students had participated.

Data were collected at the end of the experimental action by administering a questionnaire through Google Forms. On the Likert scale considered for measurement, the level of agreement ranged from "absolutely agree" to "agree," "neither agree nor disagree," "disagree," "absolutely disagree.

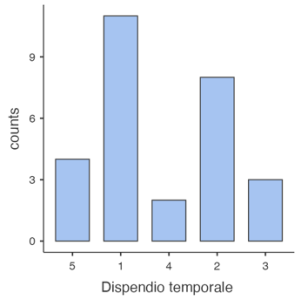
In this regard, frequency analysis revealed that 53.5 percent of the students absolutely agree that ChatGPT is a useful teaching tool (3.93. ± 1.05.), 35.7 percent disagree that it is useful, and 3.5 percent are in between (Graph 1)



Graph. 1 (students' evaluation of the chatbot's usefulness (values expressed in this section ranged from 3 to 5)

Regarding the fun component, i.e., relating to the enjoyability of the chatbot, 19.6% absolutely agree that ChatGPT was a fun tool; in fact, the analysis finds that overall, the sample ranks in the gradients of "agree," "neither agree nor disagree," and disagree with a mean of 3.35 and a standard deviation of 1.18.

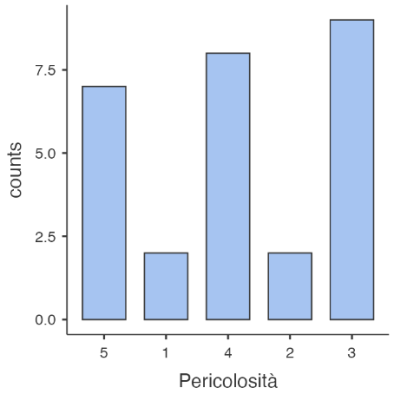
With respect, on the other hand, to the assessment of the degree of time-consumption resulting from the use of ChatGPT, 39.2 percent absolutely disagree that ChatGPT is a time-consuming tool, when in contrast 14.2 percent are in the mirror position in that they absolutely agree that the chatbot is time-consuming (2.39 ± .1.60) (Graph 2).



Graph. 2 (Assessment of time expending)

Regarding the students' assessment of their perception of the innovativeness of this technology, 67.8 percent strongly agreed that the chatbot is innovative, 28 percent were in simple agreement, and 3.5 percent were in an intermediate degree ("neither agree nor disagree), with a mean of 3.71 and a d.s. of 1.07.

The perceived challenge inherent in using ChatGPT as an instructional support, as found by the analysis, was high for 32.14% of the subjects who see the chatbot as a tool full of challenges. On the other hand, 14.2% agreed that the chatbot is challenging, albeit not absolutely, and only 7.15% of the students strongly disagreed. Finally, it is significant to note that the analysis of the sample's perception of the risks inherent in the use of ChatGPT sees only a low percentage of students (3.6429±1.1528), precisely 25%, in complete agreement with the statement "ChatGPT is dangerous in terms of the veracity of information" (Graph 3)



Graph. 3 (students' assessment of chatbot dangerousness).

3.2 Qualitative Analysis.

After the laboratory activity conducted, students were asked to take stock of the experience in terms of strengths and weaknesses on the use of Chat GPT.

The data were analyzed according to a thematic analysis that involved the development of several steps (Braun & Clarke, 2006): familiarization with the data (1), generation of initial codes (2), organization of the codes by theme (3), review of the themes with respect to the first level (the extracts) and the second level (the entire data matrix) (4), definition of the themes (5), and production of the report (6).

Following this approach, the themes found on the strengths indicated by the students are two: Utilization and Output. There are a total of 11 codes filed within the two thematic areas. With respect to the theme "Usage" the 5 codes identified are: speed, immediacy/quickness, ease, intuitiveness, and ability to file answers. With respect to the theme "Output" the 6 codes identified are: accuracy, clarity, consistency/pertinence, reliability, conciseness, fluency.

Thematic analysis of identified strengths		
Theme	**Codes**	**Extracts**
Use	Speed	"Speed in formulating concepts." "Speed in finding information."
	immediacy/rapidity	"Immediate answers for full-text writing."

		"Quick Search.”
	ease	"Ease of Use”
	Intuitiveness	"Clear and intuitive mode of use"
	ability to archive responses	"Chat archive so you can resume the search at any time."
Output	accuracy	"Comprehensive explanation of content."
	clarity	"Through stimulus we got clear definitions" "The comprehensive explanation"
	consistency/pertinence	"Ability to process and generate text consistently."
	reliability	"The relevance/truthfulness of the content of the chosen area" "Ability to formulate a coherent and informative text by providing quick and accurate answers on the topic posed."
	concision	"He responds immediately, clearly and straightforwardly."
	fluidity	"Fluidity of information". "Immediacy in transporting the content made between various digital media."

Table 3 (thematic analysis strengths)

There are four themes found on the weaknesses indicated by students: intelligibility, sources, usage, and output. There are a total of 11 codes filed within the four theme areas. With respect to the theme "Usage" the identified code is: absence of guidance. With respect to the theme "Output" the 2 identified codes are: generalization, limited knowledge. With respect to the theme "Sources" the 3 identified codes are: uncertainty of information, irrelevance of answers, non-explication of sources. With respect to the theme "Intelligibility" the 4 codes identified are: presence of lexical repetition, non-contextualization, inability to draw from experience, reflexive inability.

Thematic analysis of identified weaknesses		
Theme	**Codes**	**Extracts**
Use	absence of guidance	"Little Usage Information"
Output	generalization	"Tends to overgeneralize content." "Generality about the information received"
	limited knowledge	"postpones limited knowledge."
Sources	uncertainty of information	"Reliability to be ascertained." "Risk of uncertain data whose veracity must be verified."
	irrelevance of answers	"Answers not related to the question"
	non-explication of sources	"Takes information without making sources explicit." "Lack of sources."
Intelligence	presence of lexical repetitions	"Syntactic errors in the text such as repetitions"
	non-contextualization	"Absence of the dimension of specific context"
	Inability to draw from experience	"Inability to provide answers based on personal experience."
	reflexive inability	"Is unable to create reflective thinking is contextualized."

Table 4 (Thematic analysis weaknesses)

Conclusions

Considering the results of the quali-quantitative analysis illustrated above, the quantitative data regarding the criteria contained in the meta-reflection form indicate that the students in the sample generally followed the directions given and took into consideration the suggestions in the dedicated section of the form aimed at encouraging critical and responsible use of ChatGPT.

In particular, the only areas in which students would seem to have found some obstacle in following the recommendations are those related to the appropriateness of entering fewer than 4,000 words, providing examples, and taking an unreasonable amount of time to formulate the prompt.

Therefore, it can be concluded that, on the one hand, the evaluation tool of the meta-reflection sheet seems to be an effective tool in accompanying and activating for the 160 students a metacognitive and reflective path on the use of AI software in the educational field.

On the other hand, the detected areas of criticality suggest the need for further attention, within the framework of scientific reflection on the topic, to the mechanisms of chatbot use and the consequent elaboration processes enacted by the students, especially about the identified thematic spheres.

The conclusions that can be drawn pertaining to the questionnaire administered at the end of the experimental action also show that students' attitudes toward ChatGPT use in educational contexts - even following a targeted educational intervention - result in feelings of optimism, trust, enjoyment, and familiarity.

References

Albanese, M. & Compagno, G. (2022). *La valutazione di attività neurodidattiche. Fondamenti, tecniche e strumenti*. Roma: Anicia Editore.

Albanese, M. & Compagno, G. (2023). The neuroeducation perspective in the evaluation of the teacher's methodological-didactic skills. *Giornale Italiano Di Educazione Alla Salute, Sport E Didattica Inclusiva*, 1.

Banks, D. C. (2023). *ChatGPT caught NYC schools off guard. Now, we're determined to embrace its potential*. Chalkbeat.

Benzato, M. & Minello, R. (2002). *Imparare insieme. Laboratorio di didattica dell'apprendimento cooperativo.* Armando Editore.

Berthier, J-L., Borst, G., Desnos, M., Guilleray, F. (2018). Les neurosciences cognitives dans la classe. Guide pour expérimenter et adapter ses pratiques pédagogiques. Paris: ESF Sciences Humaines.

Buchberger, B. *Is ChatGPT Smarter than Master's Applicants?* Research Institute for Symbolic Computation: Linz, Austria, 2023.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Chen, L., Chen, P., &Lin, Z. (2020). Artificial intelligence in education: A review.Ieee Access,8, 75264–75278

Damiani, P., Santaniello, A., Gomez Paloma, F. (2015). Ripensare la Didattica alla luce delle Neuroscienze: Corpo, abilità visuospaziali ed empatia: una ricerca esplorativa. Giornale Italiano della Ricerca Educativa – Italian Journal of Educational Research, VIII, 14.

Drivas, T., & Doukakis, S. (2022, September). Introducing the Fundamentals of Artificial Intelligence to K-12 Classrooms According to Educational Neuroscience Principles. In 2022 7th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM) (pp. 1-7). IEEE.

Erfiani, Y.P.F., Neno, H. (2018), The Effect of Jigsaw Method to Improve EFL Students' Vocabulary Ability, *Metathesis: Journal of English Language Literature and Teaching,* Vol. 2, No. 2, pp. 171-183.

Fuchs, K. Exploring the opportunities and challenges of NLP models in higher education: Is Chat GPT a blessing or a curse? *Front. Educ.* 2023, *8*, 1166682.

Guillén, J. C. (2021). Neuroeducazione in classe. *Dalla teoria alla pratica, Roma: Il Bruco farfalla*.

Halaweh, M. ChatGPT in education: Strategies for responsible implementation. *Contemp. Educ. Technol.* 2023, *15*, 421

Huang, A. Y., Lu, O. H., & Yang, S. J. (2023). Effects of artificial Intelligence–Enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education*, *194*, 104684.

Javaid, M.; Haleem, A.; Singh, R.P.; Kahn, S.; Khan, I.H. Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Trans. Benchmarks Stand. Eval.* 2023, *3*, 100–115

Lo, Chung Kwan. (2023) "What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature" *Education Sciences* 13, no. 4: 410. https://doi.org/10.3390/educsci13040410

Mattingly, R.M., VanSickle, R.L. (1991). *Cooperative Learning and Achievement in Social Studies: Jigsaw II,* Institute of Education Sciences, https://eric.ed.gov/?id=ED348267 **,** pp.20.

Minello, R. (2020). Neuroscienze cognitive in aula: le condizioni d'uso. *Formazione & insegnamento*, *18*(4), 26-34.

Minello, R., (2011). Polisemie. In Minello, R., Margiotta, U., POIEIN. La pedagogia e le scienze della formazione (103-188). Lecce, Pensa MultiMedia.

Montenegro-Rueda, M.; Fernández-Cerero, J.; Fernández-Batanero, J.M.; López-Meneses, E. Impact of the Implementation of ChatGPT in Education: A Systematic Review. *Computers* 2023, *12*, 153. https://doi.org/10.3390/computers12080153

Oranga, J. (2023). Benefits of Artificial Intelligence (ChatGPT) in Education and Learning: Is Chat GPT Helpful? International Review of Practical Innovation, Technology and Green Energy (IRPITAGE), 3(3), 46–50. https://doi.org/10.54443/irpitage.v3i3.1250

Slavin, R. E. (1991). *Student team learning: A practical guide to cooperative learning*. National Education Association Professional Library, PO Box 509, West Haven, CT 06516 (Stock No. 1845-1-00, $18.95).

Strzelecki, A. To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interact. Learn. Environ.* 2023, 1–14.

Susnjak, T. (2022). ChatGPT: The end of online exam integrity?. *arXiv preprint arXiv:2212.09292*

Wang, T.; Lund, B.D.; Marengo, A.; Pagano, A.; Mannuru, N.R.; Teel, Z.A.; Pange, J. Exploring the Potential Impact of Artificial Intelligence (AI) on International Students in Higher Education: Generative AI, Chatbots, Analytics, and International Student Success. Appl. Sci. 2023, 13, 6716.

Vinci, V. (2019). Neuroscienze e agire educativo. Nuova Didattica, reperibile presso http://nuovadidattica.lascuolaconvoi.it/agire-educativo/la-mediazione-plurale-nel-lavoro-educativo/neuroscienze-e-agire-educativo.

Zhai, Xiaoming, ChatGPT User Experience: Implications for Education (December 27, 2022).

.