Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno5 n.4 (2025)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

Transitional Inferences: Cognitive and Technological Hybridizations for an Epistemically Responsible Education

Antonella Nuzzaci [1]

[1] University of Messina; antonella.nuzzaci@unime.it

Abstract: In today's educational landscape, the growing presence of Artificial Intelligence (AI) is reshaping learning processes and introducing new epistemological challenges. Inferential practices – deductive, inductive and above all abductive – today constitute a meeting point between the human mind and algorithmic logic. The present paper explores these "fertile hybridizations" between cognitive models and intelligent systems, analyzing how symbolic, sub-symbolic and neuro-symbolic technologies implement the different forms of inference and how these influence educational contexts. The research is developed in three phases: (1) a conceptual reconstruction of the inferential forms in the logical tradition and in their algorithmic translation; (2) a critical analysis of AI-based educational technologies, with a focus on intelligent tutoring, adaptive learning and predictive assessment tools; (3) the proposal of a teaching model based on hybrid inference and oriented towards critical digital citizenship. The expected results concern the development of transversal epistemic skills such as critical awareness of sources, uncertainty management and the ability to distinguish between intuitive and explainable inference. In this sense, education is not configured as a simple adoption of technologies, but as the construction of a dialogical, cooperative and ethically founded environment, capable of forming responsible citizens in the age of AI.

Keywords: Hybrid inference, Explainable Artificial Intelligence (XAI), Epistemic responsibility, Critical abduction, Digital citizenship, Artificial intelligence in education

1. Introduction

In today's educational context, the growing presence of Artificial Intelligence (AI) is profoundly reshaping the ways in which knowledge is acquired, interpreted, and produced. AI is no longer confined to the mediation of content; it actively participates in inferential processes, transforming modes of thought, decision-making processes, and systems of meaning-making (Floridi, 2023). In light of these transformations, a critical reflection is needed on the "fertile hybridizations" that emerge between human cognitive models, algorithmic inference systems and educational environments. This paper aims to explore these hybridizations, assuming inference – be it deductive, inductive or above all abductive – as a privileged space of interaction between mind, machine and pedagogy. It is not only a matter of investigating, on a theoretical level, the functioning of inferential models (epistemological perspective), but also of understanding how they concretely affect educational processes, from teaching to learning (pedagogical perspective): the path starts with a reconstruction of inferential models, to analyze how AI systems implement them and how these processes affect educational practices, raising new challenges related to intuition, explainability and epistemic responsibility (Peirce, 1931-1958; Kahneman, 2011). As recent studies suggest (Chen et al., 2023), intuition, when integrated with critical thinking, can function as an effective epistemic filter in interactions between humans and intelligent systems.

The research is configured as a path divided into three interconnected phases, aimed at investigating the relationship between inference, artificial intelligence and educational practices.

In the first phase, a conceptual framework of the main types of inference – deductive, inductive and abductive – is outlined with particular attention to the ways in which these processes are implemented in symbolic, sub-symbolic and neuro-symbolic artificial intelligence systems (Valiant, 2013).

The second phase adopts a critical-analytical perspective, exploring educational technologies based on artificial intelligence. In this perspective, the automatic inferential mechanisms underlying intelligent tutoring systems, adaptive learning platforms and predictive assessment tools are examined, with the aim of assessing their impact on training processes.

Finally, the third phase assumes a design orientation, proposing a teaching model based on mixed inferential practices. This model aims to develop students' critical skills in interacting with AI systems, encouraging them to interrogate algorithmic outputs and recognize the conjectural and non-definitive nature of the responses produced by generative systems (Gunning et al., 2019).

In this way, the study intends not only to offer a theoretical reconstruction and a critical evaluation of inferential practices applied to AI, but also to outline operational paths for a pedagogy capable of integrating, problematizing and transforming the interaction between mind, machine and learning.

The theoretical and methodological framework draws on the epistemology of uncertainty, emphasizing the ability to navigate ambiguity and construct meaning within hybrid cognitive environments. The main goal is to develop a teaching model that promotes hybrid inference as a transversal and foundational skill for digital citizenship. The model integrates critical abduction practices, simulations in AI-based environments, algorithmic explanation activities (Explainable AI) and reflection on decision-making processes.

Recent studies suggest that intuition, when paired with critical reasoning, can serve as an effective epistemic filter in human-AI interactions (Chen et al., 2023). This model is part of a pedagogical paradigm that does not oppose human and artificial intelligence, but integrates them in a dialogical, cooperative and ethically based learning context. This approach presents itself as a concrete example of "fertile hybridization", capable of addressing the cognitive and ethical challenges posed by algorithmic society.

2. Theoretical framework

The growing diffusion of Artificial Intelligence (AI) in different social, professional and cultural contexts marks a radical transformation of cognitive and educational processes. These are no longer just technical support tools, but systems that actively participate in inferential processes, changing learning methods, decision-making logics and forms of meaning construction (Floridi, 2023). AI does not limit itself to "calculating": it infers, producing deductive, inductive and abductive conclusions that condition classifications, recommendations and interactions (Johnson-Laird & Byrne, 1991; Peirce, 1931–1958). The theoretical framework linking inference and artificial intelligence (AI) is based on the classical distinction between deduction, induction, and abduction, as outlined in Peirce's (1931–1958) philosophical-logical tradition. Deduction ensures the truth of conclusions when it derives from premises valid according to formal rules; induction, on the other hand, generalizes from particular observations, formulating probabilistic laws; finally, abduction proposes plausible explanatory hypotheses, without guaranteeing absolute certainty. In AI systems, these inferential modalities find different implementations, namely: symbolic systems are mainly based on deduction, exploiting logical rules and structured knowledge bases; sub-symbolic systems, such as deep neural networks, operate primarily by induction, learning patterns and correlations from data; neuro-symbolic systems try to combine these strategies, integrating deductive and inductive capacities and thus paving the way for computational forms of abduction, capable of generating plausible hypotheses and intelligible explanations (Valiant, 2013).

This technological landscape has significant implications also in education, where uncertainty, far from being a mere limitation, becomes an epistemic resource, since the ability to assess plausibility, integrate intuition and critical reasoning and navigate ambiguity is central to contemporary learning (Kahneman, 2011). In this context, Explainable AI (XAI) tools are particularly relevant, as they allow to make the decision-making processes of intelligent systems transparent and to promote teaching practices oriented towards explanation, critical reflection and argumentation (Gunning et al., 2019). The synergy between inferential forms of AI and uncertainty-based pedagogical approaches therefore represents a fertile ground for developing advanced cognitive skills, which combine logical reasoning, intuition and the ability to generate and evaluate hypotheses in complex contexts.

A further point of connection concerns the relationship between inferential practices and critical digital literacy. If digital citizenship requires the ability to navigate complex information environments, filter sources, and recognize algorithmic biases, then educating in inference becomes an integral part of this skill. In particular, distinguishing between deductive, inductive, and abductive inferences allows students to better understand not only how AI processes data, but also how digital discourse and online information construct meanings and guide decisions. Critical digital literacy, in this sense, is not limited to technical literacy, but is configured as an epistemic practice: teaching to evaluate the plausibility of an output, to problematize the opacity of a sub-symbolic model or to recognize the limits of algorithmic generalizations means developing forms of conscious citizenship. Inference, therefore, is not just an internal cognitive process, but a social and critical skill that allows one to exercise responsibility in the contemporary digital ecosystem.

In the field of education, this scenario opens up a twofold front. On the one hand, smart technologies make it possible to expand the possibilities of personalized learning through adaptive platforms, automatic tutoring systems and predictive assessment tools (Ane & Nepa, 2024; Luckin, 2018). On the other hand, they raise new epistemic and ethical challenges, where the "inferential opacity" typical of sub-symbolic and generative models, capable of providing plausible results without making explicit the criteria followed, risks reducing the student to a simple passive user (Pedwell, 2023; Suresh & Guttag, 2021).

In this framework, educating to inference becomes a transversal and priority objective in teaching-learning processes. Every educational activity—from solving a mathematical problem to understanding a literary text, from scientific planning to historical interpretation—is based on inferential processes, that is, on the ability to move from the known to the unknown by constructing hypotheses, generalizations, and explanations (Bruner, 1961; Vygotsky, 1978).

AI, as an "epistemic partner", requires us to rethink these dynamics, integrating into the training course not only the knowledge of disciplinary contents, but also the ability to interrogate and critically evaluate the inferences produced by machines (UNESCO, 2019; Holmes, Bialik & Fadel, 2019). Other authors have elaborated similar concepts, talking about *epistemic quasi-partnerships* (Dorsch & Moll, 2024) or *symbiotic learning partnerships* (Zhang, 2024), highlighting the need for conscious collaboration between human and artificial agents.

A central element of this transformation is the role of intuition. As Kahneman (2011) points out, human thought operates through two systems: a fast, intuitive, and heuristic one (System 1) and a slow, reflective, and logical one (System 2). Educating to inference in a hybrid context means stimulating both, avoiding both fideism towards algorithmic responses and the a priori rejection of their potential. Recent studies (Chen et al., 2023) show that intuition, when supported by critical reflection and metacognition, can function as an effective epistemic filter to assess the reliability of AI-produced responses. In this sense, Polanyi's (1966) observation that "we know more than we can say" acquires new relevance today, as education is called upon to form individuals capable of valuing tacit knowledge without sacrificing argumentative rigor and transparency.

In light of these considerations, the introduction of AI in educational contexts cannot be understood as a simple technological advance, but requires a profound redefinition of new training paradigms. It is necessary to move from a teaching centered on the transmission of content to a pedagogy of mixed inference, which enhances the interaction between human reasoning and algorithmic logic, stimulates the exploration of uncertain scenarios and cultivates the epistemic responsibility of students (Nuzzaci, 2025). As Eco (1979) suggests, every interpretative process is an "inferential walk" and, in the age of AI, schools and universities have the task of teaching how to consciously walk such paths, continuously negotiating between human hypotheses and artificial explanations.

3. The forms of inference in the philosophical logical tradition

Inferential thinking forms the backbone of human cognitive processes and, increasingly, of artificial systems (Flach, 2012). The philosophical tradition distinguishes three fundamental forms of inference, which over time have constituted not only tools of knowledge, but also models for the development of artificial intelligence and pedagogical reflection. Before analyzing their educational and technological implications, it is appropriate to briefly dwell on the three types of inference – deduction, induction and abduction – clarifying their distinctive characteristics and the role they play both in human thought and in teaching and artificial intelligence systems.

Deduction

Deduction ensures the validity of a conclusion when its premises are true. It is the heart of formal logic and represents the paradigm of epistemic certainty. Aristotle encodes it in syllogisms (Aristotle, trans. 1984), while in the twentieth century it finds development in the axiomatic systems of mathematical logic (Hilbert & Ackermann, 1959; Gödel, 1931/1992). On the educational level, deduction is reflected in the teaching of logical and argumentative thinking, which represents a fundamental skill for critical training and for the elaboration of well-founded judgments (Lipman, 2003). In artificial intelligence, deduction underpins symbolic systems and automated formal logics, such as inference engines in expert systems, which apply rules to knowledge bases to produce rigorously valid conclusions.

Induction

Induction proceeds from the particular to the universal, formulating probabilistic generalizations. Although it does not guarantee definitive conclusions, it is fundamental for the construction of scientific laws and for the empirical validation of hypotheses (Bacon, 1620/2000; Mill, 1882). It constitutes the pivotal method of modern experimental science and remains central even in contemporary empirical research. In educational contexts, induction translates into teaching methodologies based on experience and discovery (Bruner, 1960/2009), in which the student is guided to infer general rules from concrete cases and real situations, developing generalization and adaptation skills. In AI, induction is the founding principle of machine learning systems, which learn from large amounts of data to extract patterns, predict future events, and adapt to new situations.

Abduction

A concept introduced and systematized by Charles Sanders Peirce (1931–1958/1998), abduction consists in the formulation of explanatory hypotheses in the face of incomplete or unexpected observations. It is conjectural, open-ended, and constitutes a form of creative reasoning, closely linked to the process of scientific discovery and intuition (Magnani, 2009). Abduction plays an increasing role in educational processes, as it stimulates creativity, problem-solving skills and competence in dealing with uncertainty. In the age of artificial intelligence, abduction is particularly relevant in the face of generative systems based on neural networks, which produce plausible but not definitive answers: educating to their interpretation means developing students' ability to formulate and evaluate hypotheses, even in the absence of complete information (Scardamalia & Bereiter, 2014).

These three inferential modalities should not be considered as watertight compartments, but as interconnected and often overlapping functions within human thought. The most advanced teaching is not limited to transmitting knowledge, but promotes inferential competence as a transversal skill, capable of integrating logical certainty, empirical probability and hypothetical creativity. In this sense, schools and universities today have the responsibility of preparing individuals capable of consciously moving between deduction, induction and abduction, developing critical, flexible thinking capable of dialoguing with new forms of artificial intelligence.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

4. Inferential implementations in Artificial Intelligence systems

With the advent of Artificial Intelligence (AI), inferential practices have taken on a computational dimension, transforming into operational models capable of processing, learning and generating knowledge in an increasingly autonomous way (Russell & Norvig, 2021). We can distinguish three large families of approaches, each of which favors a specific form of inference, albeit with inevitable overlaps and reciprocal contaminations:

1) symbolic AI: born in the 50s and 60s, it is based on the explicit representation of knowledge through rules, logics and ontologies. It reproduces deductive reasoning in particular, applying chains of rules to derive conclusions from codified premises (Newell & Simon, 1976). The expert systems of the 80s, such as MYCIN in the medical field, are a paradigmatic example (Shortliffe, 1976). Symbolic AI still retains relevance in fields that require transparency and explainability, such as law, clinical diagnostics, and knowledge engineering. On the educational level, this approach allows the development of teaching practices oriented towards formal logic and rigorous reasoning, offering students transparent models with which to confront and understand the coherence between premises and conclusions.
2) sub-symbolic AI: characterized by the use of artificial neural networks and machine learning techniques, it does not use explicit rules but mathematical-statistical models capable of learning from data. An inductive approach prevails here, which generalizes from examples and observations to produce predictions or classifications (Goodfellow, Bengio, & Courville, 2016). The opacity of such systems, often defined as black box, however, makes it difficult to explain the results, with significant ethical and pedagogical implications: teaching how to interpret and problematize the output of such systems becomes an integral part of critical training (Burrell, 2016). In education, this means getting students used to working with probabilistic models, recognizing the provisional nature of generalizations and learning to balance trust and skepticism towards algorithmic predictions.
3) neuro-symbolic AI: more recent, aims to combine the advantages of symbolic formalization with the predictive power of the sub-symbolic (Garcez, Besold, De Raedt, Földiak, Hitzler, Icard, Kühnberger, Lamb, Miikkulainen, & Silver, 2019). This paradigm opens up new perspectives for the implementation of computational abductive processes, capable of formulating plausible hypotheses starting from incomplete or uncertain information (Valiant, 2013). In this way, AI does not limit itself to classifying or predicting, but becomes capable of proposing explanations, approaching creative and hypothetical forms of reasoning. From an educational point of view, this opens up innovative scenarios to stimulate the ability to elaborate hypotheses, to deal with uncertainty and to use error as a learning opportunity, training a conjectural and flexible form of thinking.

From this it is clear that the current transformation does not only concern the quantity of data processed, but the very quality of the inferential processes. This change has profound consequences both on the cognitive level, as it redefines the ways in which human beings interact with knowledge, and on the educational level, where it becomes crucial to develop critical interpretation skills, understanding the limits of different inferential models and integrating human reasoning and artificial systems (Holmes, Bialik, & Fadel, 2019). The contemporary educational challenge is therefore to prepare individuals capable not only of using AI, but also of understanding its epistemological assumptions and ethical implications, thus becoming aware citizens in the digital age.

5. Epistemology of uncertainty and inferential processes

The interaction between students and intelligent systems raises crucial epistemological questions. Interaction with AI-based systems brings with it an increase in cognitive uncertainty: students and teachers are confronted with outputs that rarely have absolute validity, but which require to be interpreted, justified and, sometimes, reworked. In this context, uncertainty should not be understood as a limit, but as a generative condition of learning, as the ability to move between ambiguities, evaluate the plausibility of answers and integrate intuition and critical reasoning is today a fundamental epistemic skill (Kahneman, 2011). Explainable AI (XAI) offers tools to make algorithmic processes more transparent and to promote teaching that values explanation, argumentation and critical reflection (Gunning et al., 2019).

In this framework, contemporary epistemology, and particularly the theories of Kahneman (2011), emphasizes the need to manage the tension between two complementary cognitive modalities: System 1, rapid, intuitive and automatic, often associated with abductive inferences; and System 2, slow, analytical and reflective, typical of deduction and critical control. Teaching, therefore, must aim to develop in students the ability to activate and balance both systems, exploiting creative intuition without sacrificing analytical verification, especially in interaction with complex AI systems. For this reason, AI-mediated teaching requires the development of an epistemic competence capable of integrating the two systems, avoiding both fideism towards algorithmic responses and the a priori rejection of their potential.

One of the most critical aspects of sub-symbolic systems is their opacity; in fact, users receive answers and predictions without being able to reconstruct the logical steps that generated them (Burrell, 2016). Translated into the educational field, this phenomenon risks compromising trust and reducing learning to a mere passive reception, where the student tends to accept the output of the machine without developing awareness of the process that produced it. The result is a potential erosion of critical trust, an essential element for the formation of autonomous and responsible citizens.

The *Explainable AI* (XAI) paradigm (Gunning et al., 2019) is located exactly at this intersection, offering an alternative approach that aims to make algorithmic inferences transparent through the explanation of the criteria and data underlying a given conclusion. In this sense, XAI is not only a technical challenge for developers, but also a pedagogical challenge, which introduces the possibility of transforming the relationship with AI into an opportunity for active learning.

In this sense, the didactic use of XAI can promote practices of critical analysis, comparison, and argumentation, encouraging students to interrogate, evaluate, and even challenge the results produced by intelligent systems. The goal is not to replace human judgment with algorithmic judgment, but to develop the ability to dialogue with it in a conscious way, activating processes of co-construction of knowledge.

6. Towards a hybrid inference model

The introduction of Explainable AI (XAI) opens the way to a broader perspective, where the convergence between human cognitive models and algorithmic inferential processes generates a new epistemic scenario, which we can define as fertile hybridization. In other words, XAI is not limited to making an algorithm work or producing a prediction, but aims to make the decision-making processes of AI systems understandable and transparent. This approach is critical, as many modern systems, such as deep neural networks, operate as "black boxes," producing accurate results, but the internal logic that generates them often remains opaque. XAI, therefore, makes it possible to understand how and why AI reaches certain conclusions, fostering a deeper and more critical knowledge of intelligent systems.
The main objectives of XAI are:

1. Transparency → clearly show how the system processes information.
2. Explainability → allow users to understand the reasons for an automated decision.
3. Reliability and trust → help people and professionals to use AI with awareness, especially in sensitive areas such as medicine, law or education.

To exemplify, it is possible to give an example. If an AI system suggests a grade to a student, the XAI makes it possible to understand what data and criteria led to that result, so the teacher or student can evaluate, discuss or correct the decision.
In summary, XAI transforms AI from a "black box" to a teaching tool suitable for facilitating critical learning and supporting trust in the use of intelligent systems.
It is clear, therefore, that hybridization does not imply the replacement of human rationality with artificial rationality, but rather their cooperative integration, in which each system makes its specific potential available.
In this context, critical abduction emerges as a key competence. Students must learn to recognize the conjectural character of algorithmic responses, to verify their plausibility and to use them as a stimulus for further hypotheses and reflections. Abduction thus becomes the privileged meeting place between human intuition and the generative power of AI, promoting new forms of educational thinking in which interaction with machines does not weaken, but becomes an opportunity to strengthen the ability to think critically, argue and create knowledge.

7. Methodology

The present research, currently underway, is developed in three main phases, integrated with each other and oriented towards the exploration of inferential practices in the age of artificial intelligence.

The research is structured as a qualitative, design-based study, characterized by a cyclical progression of analysis, design and experimentation. The goal is not only to describe phenomena, but to develop and test an innovative teaching model, capable of being progressively improved through feedback and critical reflection. In this phase, the study takes on the features of an exploratory pilot, useful for assessing the transferability of the model and collecting preliminary evidence on the epistemic skills developed by the students.

The first phase of *conceptual analysis* involves the reconstruction of classical inferential models – deduction, induction and abduction – and the analysis of their use and implementation in symbolic, sub-symbolic and neuro-symbolic artificial intelligence systems (Valiant, 2013). The goal is to understand the theoretical and operational underpinnings of the different inferential modalities, both in human thinking and in AI.

The second phase of *critical analysis of educational technologies* is aimed at critically analyzing AI-based educational technologies as AI-based tools, with particular attention to the automatic inferential mechanisms underlying intelligent tutoring systems, adaptive learning platforms and predictive assessment tools, with particular attention to the inferential mechanisms implicit in their algorithms. The aim is to highlight pedagogical opportunities and criticalities, emphasizing the need to develop epistemic skills, such as managing uncertainty and critically evaluating machine-generated responses.

The second phase relating to the *design proposal* is oriented towards defining a precise didactic design, which proposes a model based on hybrid inferential practices. The theoretical-methodological framework is based on the epistemology of uncertainty, understood as a perspective capable of integrating intuition and explainability in hybrid learning contexts. The model is divided into four dimensions (Table 1):

- *Critical abduction*: encouraging students to formulate interpretative hypotheses in the face of uncertain or ambiguous algorithmic outputs;
- *Simulations in AI environments:* use adaptive platforms to develop uncertainty management and decision-making skills;
- *Algorithmic explanation*: promote educational activities based on the analysis of the logic underlying the responses generated by the systems (Explainable AI);
- *Argumentative reflection:* guide students in the construction of inferential maps and in the comparison between human and artificial responses.

Table 1. Operational dimensions of the hybrid inference model and related teaching activities

Operating size	Classroom teaching activities	Tools / Resources	Educational objectives
Critical abduction	- Analysis of uncertain or ambiguous algorithmic outputs - Guided discussion on possible interpretations - Formulation of explanatory hypotheses	- Generative systems such as ChatGPT - Examples of incomplete or partial data	- Develop hypothesis skills - Training creative and conjectural thinking - Managing uncertainty in complex contexts
Simulations in AI environments	- Use of adaptive platforms or simulators for guided experiments - Troubleshooting with immediate feedback - Simulated decision-making scenarios	- Adaptive learning platforms (e.g. Smart Sparrow, Khan Academy) - Interactive AI simulations	- Develop decision-making skills - Train uncertainty management - Promote experience-based learning
Algorithmic Explanation (XAI)	- Analysis of the logic behind AI-generated responses - Identification of inferential steps - Discussion of why the system produced a certain response	- XAI tools - Generative systems with step-by-step explanations	- Understand the inferential mechanisms of the machine - Stimulate critical thinking and argumentation skills - Evaluate reliability and plausibility of answers
Argumentative reflection	- Collaborative construction of inferential maps to visualize reasoning - Comparison of human responses and AI responses - Review and justify your inferences	- Digital whiteboards/concept map software (e.g. Miro, CmapTools) - Generative AI systems	- Develop argumentation and comparison skills - Promoting epistemic awareness - Improve communication and collaboration skills

This will involve the structuring of a pilot teaching module, currently in the *try-out* phase, which includes: (a) guided analysis of responses generated by AI systems (e.g., ChatGPT), (b) collaborative construction of inferential maps, (c) comparative evaluation of human responses versus those generated by the machine, and (d) activities involving justification and argumentative review. Through these practices, the goal is to develop epistemic skills such as source awareness, critical appraisal of plausibility, uncertainty management, and the ability to distinguish between intuitive and explainable inference.

Its goal is to train students to critically interact with AI, to interrogate algorithmic outputs, and to recognize the conjectural and non-definitive nature of the responses generated by generative systems, thus promoting key epistemic skills such as source awareness, plausibility assessment, and uncertainty management (Gunning et al., 2019).

Activities can be modular and can be carried out individually or combined into a larger project.

The design involves the use of continuous feedback, both from students to the teacher, and between students themselves, to stimulate critical reflection.

The evaluation focuses not only on the final result, but on the inferential process followed, on the ability to justify choices and to integrate intuition and explanation.

8. Expected results

Through this path, we intend to develop transversal epistemic skills, such as:

- ✓ critical awareness of sources;
- ✓ ability to assess plausibility;
- ✓ uncertainty management;
- ✓ Distinction between intuitive inference and explainable inference.

Table 2. Expected epistemic skills and survey tools in the teaching model

Skills	Indicator	Example of behavior
Critical awareness of sources	Recognizes and reports the origin of data or information	The student verifies the authority of the source of an AI prediction before using it
Ability to assess plausibility	Compare the AI response with prior knowledge or available evidence	The learner identifies discrepancies between the AI's response and known empirical data or logical rules
Uncertainty management	Can distinguish between certain and hypothetical results, weighing the degree of reliability	The student reports when an AI response is just a possibility or requires further verification
Distinguishing Between Intuitive Inference and Explainable Inference	Recognizes whether a conclusion comes from human reasoning, intuition, or an explainable model	The student clearly describes when they follow their intuition vs when they use an explanation provided by

Five main dimensions constitute common indicators of students' epistemic and cognitive development:

- ✓ Critical awareness of sources: represents the ability to select, evaluate and compare information from different sources, distinguishing between reliable and unreliable data. This skill is crucial for developing autonomous thinking and for knowing how to correctly integrate human and algorithmic inputs;

- ✓ uncertainty management: indicates the ability to recognize ambiguous or incomplete situations, formulate hypotheses in uncertain contexts and modulate the level of confidence of one's inferences. This element reflects the epistemic maturity and cognitive flexibility of the students;
- ✓ Explainable and argued inference: involves the ability to transform initial insights or hypotheses into coherent, structured, and justified explanations. It promotes the development of cognitive transparency, making students' thinking communicable and verifiable, both among peers and with respect to AI tools;
- ✓ human-AI critical integration: refers to the competence to interact with intelligent systems in a conscious way, evaluating, contesting and reprocessing the outputs generated by AI. It allows you to develop a balanced relationship between human intuition and algorithmic reasoning, avoiding both uncritical acceptance and total rejection of technologies;
- ✓ Epistemic collaboration: reflects the ability to co-construct knowledge through dialogue and negotiation with other students and with intelligent systems. It includes sharing inferences, critical discussion, and the ability to reach reasoned conclusions in collaborative settings.

In addition to the skills already outlined (critical awareness of sources, evaluation of plausibility, uncertainty management, distinction between intuitive and explainable inference), the project involves the use of specific evaluation tools to measure the effectiveness of the model:

- Inferential skills analysis rubrics to observe how students move from intuitive hypotheses to reasoned explanations, distinguishing between deduction, induction and critical abduction;
- tools on AI perception to detect attitudes, trust and epistemic awareness with respect to the use of AI in educational contexts;
- observation grids of collaborative activities to analyze the level of interaction, negotiation and critical confrontation between students and between students and intelligent systems.
- benchmarking of outputs to compare solutions produced by students and AI systems, assessing the extent to which students are able to integrate, contest or rework algorithmic responses.

These tools do not only aim to measure performance, but to monitor the development of a conscious epistemic citizenship. The use of rubrics, questionnaires and structured observations allows to detect the degree of critical autonomy of the students, the ability to problematize the inferential opacity of sub-symbolic systems and to integrate intuition with argumentative reasoning.

The use of a triangulation between quantitative (questionnaires) and qualitative (grids, rubrics) tools offers a more complete picture, allowing us to validate the proposed pedagogical model and to identify improvement strategies.

Despite the different tools used — rubrics for the analysis of inferential skills, self-assessment questionnaires and observation grids of collaborative activities — the five dimensions are not limited to measuring traditional academic performance but offer cross-cutting metarubrics that allow to assess the development of integrated cognitive, ethical and epistemic skills. Their observation and measurement provide a more complete picture of learning, promoting conscious and critical digital citizenship in the age of AI.

Table 3. *Dimensions and tools for detecting inferential and collaborative skills*

Size / Indicator	Inferential rubric	Self-assessment questionnaire	Collaborative observation grid	Specific expertise detected
Critical awareness of sources	Evaluation of the selection and plausibility of sources used in inferences	Self-assessment of the ability to distinguish reliable and unreliable sources	Observation of discussions and comparison of sources between students	Critical thinking, evaluation of sources
Uncertainty management	Tracing the degree of certainty of the hypotheses formulated and distinguishing between certainty and conjecture	Perception of security in one's responses to incomplete data	Negotiation and discussion between students in ambiguous contexts	Cognitive flexibility, adaptation to incomplete information
Explainable and Argued Inference	Ability to transform insights into structured and justified explanations	Self-assessment of the clarity and consistency of one's arguments	Observation of how students explain reasoning and justifications	Cognitive transparency, ability to communicate inferences
Human-AI Critical Integration	Assessment of the ability to reprocess or challenge AI-generated outputs	Perception of trust and critical awareness in the use of AI	Discussion and negotiation in the integration of AI suggestions	Digital critical thinking, conscious interaction with intelligent systems
Epistemic collaboration	Analysis of how individual inferences are shared and negotiated	Self-assessment of participation and quality of interaction	Direct observation of dialogues, critical exchanges and shared construction of knowledge	Collaborative work, knowledge co-construction skills

A further expected result concerns the transferability of the teaching model to different educational contexts. In primary school, the practices of critical abduction and algorithmic explanation can be adapted in a playful and laboratory form, favoring curiosity and the early development of the ability to problematize technological outputs. In secondary school, the model finds application in specific disciplinary activities (science, mathematics, history, literature), stimulating the ability to connect intuitive inferences and formal explanations. In the university environment, on the other hand, the model can be used to promote advanced skills in critical analysis, epistemological reflection and ethical responsibility, preparing students to deal with complex artificial intelligence systems in their future professional contexts. This perspective of educational scalability represents a key element to validate the model not only as a pilot experiment, but as a replicable and adaptable tool at different levels and levels of education.

4. Conclusions

The proposed teaching model is not based on the opposition between human intelligence and artificial intelligence, but on their integration into a dialogical, cooperative and ethically oriented framework. In this sense, it represents a concrete example of "fertile hybridization", capable of responding to the cognitive and ethical challenges posed by the algorithmic society and of preparing students and citizens for a conscious, critical and constructive interaction with intelligent systems.

The analysis carried out shows how deduction, induction and especially abduction, traditionally studied in philosophy and pedagogy, find new forms of implementation in symbolic, sub-symbolic and neuro-symbolic systems, and how these hybridizations radically transform educational contexts. Explainable AI practices emerge as an indispensable tool to make algorithmic processes transparent, promote critical reflection and develop epistemic skills in learners oriented towards uncertainty management.

The hybrid inference model thus outlined opens the way to an education that is not limited to using digital tools, but also teaches how to interrogate, contest and reinterpret the outputs of machines, fostering the growth of a conscious epistemic citizenship. Education thus becomes a laboratory of cognitive resilience in which error, ambiguity and conjecture are not limits, but opportunities for learning and innovation.

The implications are many: on a pedagogical level, the model can be adapted to different school levels and disciplines; on a social level, it contributes to the construction of a critical digital citizenship, capable of orienting oneself in an information ecosystem dominated by artificial intelligence; on the ethical level, it recalls the need for shared responsibility in the design and use of educational technologies.

Finally, the path opened by this research suggests some future perspectives: the empirical experimentation of the model in different training contexts, the comparative evaluation of its effects on epistemic and metacognitive skills, and the exploration of how the new generations of generative and multimodal AI can further expand the possibilities of inferential learning. In this scenario, the main challenge is not only to teach how to use AI, but to teach how to think with and beyond AI, keeping critical autonomy, epistemic responsibility and ethical cooperation at the center.

References

Ane, B., & Nepa, G. (2024). Artificial intelligence prediction model for educational knowledge representation through learning performance. *Research on Education and Media, 16*(2), 1-10. doi: 10.2478/rem-2024-0011.

Aristotele. (1984). *Organon* (trad. it.). Roma-Bari: Laterza.

Bacon, F. (1620/2000). *Novum Organum*. Milano: Bompiani.

Bruner, J. (1960/2009). *The Process of Education*. Cambridge, Mass.: Harvard University Press.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31,* 21–32.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1). https://doi.org/10.1177/2053951715622512

Chen, J., Liao, Q. V., Vaughan, J. W., & Bansal, G. (2023). Understanding the Role of Human Intuition on Reliance in Human-AI Decision-Making with Explanations. *arXiv:2301.07255*. https://doi.org/10.48550/arXiv.2301.07255.

Dorsch, J., & Moll, M. (2024). *Explainable and Human-Grounded AI for Decision Support Systems: The Theory of Epistemic Quasi-Partnerships (EQP). arXiv preprint.* https://arxiv.org/abs/2409.14839

Eco, U. (1979). *Lector in fabula. La cooperazione interpretativa nei testi narrativi*. Milano: Bompiani.

Flach, P. (2012). *Machine Learning: The Art and Science of Algorithms that Make Sense of Data.* New York: Cambridge University Press.

Floridi, L. (2023). *The Ethics of Artificial Intelligence in Education.* Oxford: Oxford University Press.

Garcez, A. d.,Gori, M., Lamb, L. C., Serafini, L. Spranger, M., & Tran, S, N. (2019). Neural-symbolic computing: An effective methodology for principled integration of machine learning and reasoning. *arXiv:1905.06088*. https://doi.org/10.48550/arXiv.1905.06088

Gödel, K. (1931/1992). *On Formally Undecidable Propositions*. New York: Dover Publications.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* Cambridge, Mass.: MIT Press.

Gunning, D., Aha, D., & Hodges, J. (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. AI Magazine, 40(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable Artificial Intelligence. *Science Robotics, 4*(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Hilbert, D., & Ackermann, W. (1959). *Grundzüge der theoretischen Logik*. Berlin, Gottingen, Heidelberg: Springer-Verlag.

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning.* Boston: Center for Curriculum Redesign.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hove, Hillsdale: L. Erlbaum Associates.

Kahneman, D. (2011). *Thinking, Fast and Slow.* New York, NY: Farrar, Straus and Giroux.

Lipman, M. (2003). *Thinking in Education*. New York, NY: Cambridge University Press.

Luckin, R. (2018). *Machine Learning and Human Intelligence: The Future of Education for the 21st Century*. London: UCL IOE Press.

Magnani, L. (2009). *Abductive Cognition*. Berlin/Heidelberg: Springer.

Mill, J. S. (1882). *A system of logic, ratiocinative and inductive.* s.n: University of Toronto Press.

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM, 19*(3), 113–126. https://doi.org/10.1145/360018.36002

Nuzzaci, A. (2025). Promoting Inferential Processes in Educational Contexts in the Age of Artificial Intelligence. *Italian Journal of Health Education, Sports and Inclusive Didactics, 9*(2), 1-25. https://doi.org/10.32043/gsd.v9i2_Sup.1

Pedwell, C. (2023). Intuition as a "trained thing": sensing, thinking, and speculating in computational cultures. *Subjectivity 30*, 348–372. https://doi.org/10.1057/s41286-023-00170-x

Peirce, C. S. (1934). *Collected Papers of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press.

Polanyi, M. (2009). *The Tacit Dimension.* Chicago: The University of Chicago Press

Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Hoboken, NJ: Pearson.

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 397–417). New York: Cambridge University Press.

Shortliffe, E. H. (1976). *Computer-Based Medical Consultations*: New York, NY: MYCIN. Elsevier.

Suresh, H., & Guttag, J. (2021). A framework for understanding unintended consequences of machine learning. *arXiv:1901.10002v5*. https://arxiv.org/abs/1901.10002v5

UNESCO. (2019). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*. Paris: UNESCO.

Valiant, L. G. (2013). *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World*. New York: Basic Books.

Vygotskij, L. S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.

Zhang, J. (2024). *Strengthening Human Epistemic Agency in the Symbiotic Learning Partnership With Generative Artificial Intelligence, 54*(1), 1-11. https://www.researchgate.net/publication/391337614