WHEN ALGORITHMS LEARN TO SPEAK: DEBATE AS AFFECTIVE EDUCATION IN THE AGE OF ARTIFICIAL INTELLIGENCE

Rosa Carnevale
Università degli Studi di Bari Aldo Moro
carnevale.rosa1@gmail.com

Double Blind Peer Review

Citation
Carnevale, R. (2026). When algorithms learn to speak: Debate as affective education in the age of artificial intelligence. *Italian Journal of Health Education, Sport and Inclusive Didactics,* 10 (3).

Doi: https://doi.org/10.32043/gsd.v1i3.1747

Copyright notice:
© 2023 this is an open access, peer-reviewed article published by Open Journal System and distributed under the terms of the Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

gsdjournal.it
ISSN: 3103-5744

ABSTRACT
The expansion of generative artificial intelligence in educational and communicative contexts raises questions that go far beyond simple technological integration. When an algorithm is able to simulate empathy, sustain credible conversations and produce emotionally engaging texts, two distinct risks emerge: the standardization of language and the impoverishment of students' affective and relational skills. This contribution proposes Debate as a conscious pedagogical response to this challenge: not only a methodology for critical thinking, but also a training ground for affective digital literacy. The underlying thesis is that, the more communicative technologies become persuasive and apparently empathic, the more schools have a duty to educate students in the emotional awareness of communication.

KEYWORDS
Debate; generative artificial intelligence; affective digital literacy.

Received 2026-06-24
Accepted 2026-10-22
Published 2026-10-30

1. Introduction

There is a precise moment when a technological transformation stops being a technical fact and becomes an educational question. For generative artificial intelligence, that moment has already arrived. Today, a conversational system based on Large Language Models can write coherent argumentative texts, simulate empathic responses, formulate rebuttals and adapt its communicative register to the interlocutor with unprecedented effectiveness. Some studies show that GPT-4 can reach high levels of conversational persuasiveness, especially when it has information about the profile of the listener; others document that the responses of language models are perceived as empathic even though they imply no emotional experience or relational intentionality.

For a long time, schools have regarded linguistic and argumentative production as privileged evidence of the maturation of thought, strongly associated with students' cognitive development. Yet if a machine can produce a formally persuasive discourse, that equivalence no longer holds. We must ask what it means, for a student, to speak and argue before someone else. The question is not whether students may use AI to prepare a speech: it is what risks becoming weaker when language becomes immediately available, endlessly re-formulable and separated from the responsibility of the person who utters it.

It is within this problematic space that Debate acquires renewed relevance. Traditionally understood as a methodology for developing critical thinking and democratic citizenship, Debate can now be reread as a practice of affective digital literacy. By affective digital literacy this contribution refers to the ability to recognise, interpret, and regulate the emotional, persuasive, and relational effects produced by digital and artificially generated communication. It includes the capacity to identify the linguistic markers of simulated empathy, to distinguish emotional support from authentic relationship, to analyse persuasive strategies and implicit biases, and to assume responsibility for one's own speech in technologically mediated communicative environments.

In the Debate classroom, conflict is regulated, listening becomes a procedural condition, speech is exposed to the judgement of others, and the opponent becomes an interlocutor to be understood and refuted. In this sense, Debate works as a regulated pedagogical device that makes visible emotions, biases, communicative responsibilities, and relational

dynamics that no artificial system truly undergoes. It does not merely train students to produce better arguments; it places them in the embodied and relational conditions in which arguments must be listened to, contested, reformulated, and personally assumed.
This contribution is a systematised literature review based on thematic synthesis. It does not present an empirical study, nor does it aim to measure the effects of Debate on students' use of artificial intelligence. Rather, it has a propositional and theoretical-interpretive nature: it brings together research on generative artificial intelligence, affective and persuasive communication, educational risks linked to communicative delegation, and classroom Debate in order to construct a pedagogical framework. The central question is: what happens to the formation of speech, listening and dissent when students interact with systems capable of generating persuasive and apparently empathic language?

2. Research questions and methods

The review is guided by three research questions.
The first asks how recent literature describes the ability of generative AI to produce persuasive communication and simulate empathy.
The second investigates which educational risks emerge with regard to linguistic homogenisation, communicative delegation and the impoverishment of students' critical and relational skills.
The third asks how Debate can be interpreted as a practice of affective digital literacy, capable of educating students in listening, conflict regulation and responsibility for speech.
The three questions correspond to the three results discussed in the second part of the contribution.

3. Sources consulted and criteria for constructing the corpus

The bibliographic search was conducted by combining international databases, Italian databases and journals, open-access repositories and institutional websites. The following were consulted: Scopus, Web of Science, ERIC, Education Source, PsycINFO and PubMed, with Google Scholar used as an integrative source. For the Italian context, the Ministry of Education and Merit, INDIRE, RicercAzione, Form@re and university IRIS repositories were consulted.
The search strings that proved most effective were:

Area 1 – Generative AI, empathy and persuasion

"Large language models" AND "empathy AND " persuasion" NOT "sport"

"Conversational AI" AND "empathy" AND "students"
"ChatGPT" AND "empathy" AND "education"

Area 2 – Educational risks, language and relational skills

"AI literacy" AND "affective" AND "education"
"Digital literacy" AND "emotions" AND "education"
"AI" AND "relational skills"
"Language homogenization" AND "generative AI"

Area 3 – Debate, listening, dissent and affective literacy

"Classroom debate" AND "critical thinking" AND "students
"Debate" AND "pensiero critico" AND "scuola"
"Debate pedagogy" AND "dialogic education"
"metadialogue" AND "classroom discussion"
"Debate" AND "competenze socio-emotive"

Empirical contributions, reviews, theoretical studies and institutional documents were included when they were relevant to at least one of the three core areas: generative AI as empathic and persuasive communication; educational risks of delegation, homogenisation and reduction of critical thinking; Debate, dialogic argumentation, listening, emotions and metadialogue. Publications in English and Italian were considered, mainly between 2015 and 2026 for generative AI, and between 2000 and 2026 for Debate and educational argumentation. A few earlier references were used only as a theoretical framework, not as main studies in the synthesis.

Purely technical studies on AI with no connection to education, communication, empathy or persuasion were excluded, as were generic texts on digital innovation, popular articles without method or scientific bibliography, and studies on competitive Debate that could not be transferred to educational contexts.

The corpus actually read comprises 23 full contributions, distributed across three clusters: 7 on generative AI as empathic and persuasive communication; 8 on the educational risks of delegation and homogenisation; and 8 on Debate as dialogic argumentation and affective literacy.

The corpus was not intended to be exhaustive, but theoretically and pedagogically relevant. Contributions were selected because they addressed at least one of the three research questions and provided conceptual, empirical, or institutional evidence useful for constructing the thematic synthesis.

4. Method of synthesis

Given the heterogeneity of the contributions, a thematic rather than quantitative synthesis was adopted. The review does not measure the effect of Debate on the use of AI, nor does it establish experimental causal links. It constructs a pedagogical framework grounded in the full contributions read. The synthesis developed in three steps:
1) extraction of the recurring key concepts in the texts, including simulated empathy, cognitive delegation, metadialogue, emotional regulation and responsibility for speech, artificial persuasion;
2) grouping into six thematic categories: affective and persuasive artificial communication, educational risks of communicative delegation to AI, Debate as pedagogical regulation of conflict, listening and recognition of the other, metadialogue, emotions and bias, affective digital literacy;
3) interpretive integration in relation to the three research questions.
The final synthesis relates the three parts of the contribution, thus arriving at coherent conclusions on the topic.

5. Results of review

5.1 Generative AI as persuasive and apparently empathic communication

There is an idea, rooted in Western culture since Aristotle, that emotion and reason are distinct faculties, often in conflict. On one side there is logos, cold and calculating; on the other, pathos, warm and irrational. For centuries, rhetoric has taught how to use both as instruments of persuasion, but in the collective imagination 'true' intelligence has remained the intelligence that calculates, deduces and demonstrates. Artificial intelligence systems seemed to confirm this hierarchy: machines unbeatable at calculation, mediocre at human nuance, until the arrival of Large Language Models overturned it.
Today a system such as GPT-4 does not merely produce grammatically correct texts; it generates responses perceived as empathic, consoling and affectively attuned. A systematic review published in 2024 in the Journal of Medical Internet Research concluded that, in some standardised clinical scenarios, LLM responses were rated as more empathic than those of physicians in training. Not more empathic than

those of an expert psychotherapist, to be clear; but it is clear that AI had learned to recognise the linguistic patterns associated with empathy—soft tone, reformulation of others' emotions, open questions—and to reproduce them with a consistency that human beings struggle to match under stress.

The study by Salvi and colleagues on the conversational persuasiveness of GPT-4 placed the model in short multi-turn debates with human participants, a condition much closer to the real use of these systems than experiments based on static text. The results show that GPT-4 is particularly effective when it has information about the interlocutor's profile: persuasion no longer depends only on the logical quality of the argument, but also on the system's ability to calibrate the message to the possible cognitive vulnerabilities of the listener. In light of these studies, this phenomenon can be interpreted as a form of statistical empathy: a reproduction of the outward forms of feeling that may become, in certain contexts, difficult to distinguish from the original.

Rogiers and colleagues show that this persuasive function is not a side effect, but a structural feature of LLMs, which are able to adapt rhetorical strategy in very different domains. Hackenburg and colleagues add the dimension of scale: what in human discussion would require time, presence and reciprocity can be produced artificially, rapidly and distributed on a large scale. From a pedagogical perspective, this suggests that the educational issue is not only whether AI-generated texts are correct or useful, but how they produce effects of credibility, emotional proximity and persuasion.

Lee and colleagues explore an aspect that is particularly useful for teaching: artificial empathy is also a recognisable stylistic fact. Models build the empathic response through a series of recurring linguistic markers: formulas of emotional validation or reformulation of the user's concerns, as well as a reassuring tone and an orderly structure, repeated with a regularity that students educated to recognise them could learn to distinguish. Welivita and Pu complete the picture by showing how the system's availability is constant and its tone always calibrated—features that make it necessary to distinguish between an interlocutor that is always reassuring and an educational relationship that requires real presence, responsibility and sometimes the courage to disagree. Finally, the contribution by Savoldi and colleagues finally situate this scenario within the Italian context, documenting the growing use of generative

chatbots in sensitive areas such as emotional support and conflict management among younger users.
On the basis of this body of literature, the first result of the review can be formulated as follows: generative AI systems construct communicative environments capable of producing effects of persuasion and emotional support without this implying empathy or intentionality in the human sense of the term. This review therefore proposes to read such systems not simply as tools for producing language, but as environments that reshape the conditions under which students encounter persuasive and apparently empathic communication. It is therefore necessary to educate students to recognise how a text persuades, involves and orients, distinguishing the linguistic form of the relationship from real relational experience.

5.2 Pedagogical risks, linguistic homogenisation and communicative delegation

Before describing the risks, a premise is necessary: the literature analysed authorises neither the rejection of AI nor an undifferentiated alarm. The contributions examined recognise the opportunities that generative systems offer in education. The risks they identify do not concern the devices in themselves, but the ways in which they tend to be used in the absence of conscious pedagogical design.
The first risk is the one Gerlich analyses through the concept of cognitive offloading: the tendency to transfer to external tools tasks that would require attention and reasoning. The problem emerges when delegation no longer concerns routine operations, but high-level processes such as constructing an argument, evaluating a source or taking a position on a controversial issue. In such cases, the student does not use AI as a support for thinking, but as a substitute for thinking. The study by Lee and colleagues adds that the level of trust placed in the system is decisive. The more a student perceives AI as competent, the more he or she tends to reduce critical effort in evaluating outputs. This means that if a student perceives AI as more orderly and confident than he or she could be, the student will tend not to verify, not to rework, and not to assume responsibility for reasoning in the first person. In pedagogical terms, these findings suggest that the central question is not whether AI

should be used, but how it can be used while preserving the student's role as the conscious protagonist of their own learning.

The second risk is linguistic homogenisation. Sourati, Ziabari and Dehghani show that LLMs, trained on large dominant corpora, tend to reproduce the most widespread linguistic patterns, with an effect that involves not only the superficial style of writing, but also argumentative structures and ways of organising reasoning. Moon, Green and Kushlev confirm this with a precise empirical finding: the use of ChatGPT may improve the apparent quality of individual texts, but it reduces the collective diversity of productions. Teachers thus find themselves facing texts that are linguistically effective but difficult to distinguish in terms of voice, style and perspective. In light of these studies, linguistic homogenisation can be interpreted not merely as a stylistic effect, but as an educational risk concerning the formation of students' personal voice and argumentative agency.

As Zuboff notes, the standardization of language is not only an aesthetic problem, but above all a political problem in the broadest sense of the term. It concerns the capacity of individuals to participate in public conversation as distinct subjects, and not as interchangeable variants of the same template.

Institutional responses to the phenomenon have not been slow to emerge. The Guidelines of the Ministero dell'Istruzione e del Merito insist on the centrality of the person and on the need to preserve the formative function of school. The report by Tortuga and Yellow Tech documents that Italian students already use generative systems extensively, often before schools have defined a framework. Finally, Galli's contribution shifts attention to the level of educational responsibility. When a tool can produce explanations and arguments, the very meaning of the relationship between teacher, student and knowledge changes. This review proposes to read these institutional and theoretical contributions as converging around a common point: AI integration cannot be reduced to access, efficiency or regulation, but requires an explicit redefinition of educational responsibility.

The second result can therefore be formulated as follows: generative AI can support learning, but it may produce specific negative effects when introduced without pedagogical mediation. The deepest risk is not that students write with AI, but that they stop experiencing the effort of formulation, dissent and taking responsibility for what they say. From a pedagogical perspective, this suggests that the risk of cognitive

delegation becomes even more significant when it also becomes affective and communicative delegation: when students ask the machine not only to provide content, but also to find the right tone, manage disagreement or shape the emotional form of their speech. It is precisely at this point that Debate acquires a specific pedagogical function.

5.3 Il Debate as a practise of affective digital literacy

The most fruitful theoretical starting point is the distinction developed by Rapanta and Macagno between dialogue and metadialogue. Dialogue is the exchange of content; metadialogue is the capacity to reflect on how one is dialoguing, making visible the strategies used, the emotions that guide argumentative choices and the assumptions that remain implicit. Metadialogue does not emerge spontaneously; it requires explicit didactic design, and this is precisely what Debate, in its more mature forms, achieves. The educational question regarding AI is therefore translated into a series of questions that Debate makes systematic. What thesis has been constructed, and why? What tone does the text adopt, and what effect does it produce on the interlocutor?

In an earlier contribution, Rapanta interprets argumentation as critically oriented pedagogical dialogue. Critical thinking does not emerge in the well-structured monologue, but in the confrontation with someone who thinks differently and asks for reasons. Argumentative competence does not develop in the solitude of the blank page, but in the encounter with a real interlocutor who tests one's positions, an encounter that Debate structures in a systematic and repeatable way. In light of this literature, Debate can be interpreted not simply as a method for producing arguments, but as a pedagogical environment in which arguments are exposed to the pressure of listening, objection and response.

Nussbaum adds an interesting consideration: conflict is not an obstacle to learning; on the contrary, it becomes one of its conditions. Students learn more deeply when they are placed in contact with divergent perspectives and are forced to engage with that tension, provided that it is elaborative conflict, in which positions are argued and justified rather than merely opposed. This is confirmed by the studies of Asterhan and colleagues who place Debate within the tradition of dialogic pedagogy and regard it as an effective formalised version: Debate not

only encourages open participation, but also organises roles, turns, times and opposing positions so that regulated structure does not impoverish dialogue, but makes it productive. From a pedagogical perspective, this suggests that the educational value of Debate lies not in conflict as such, but in the possibility of transforming conflict into a structured occasion for reasoning, listening and relational regulation.

The affective dimension is the one most often underestimated by the literature on critical thinking, and yet it is central to this contribution. Wilkie and Ayalon show that tension, frustration, uncertainty and satisfaction are not noise that disturbs the cognitive process, but a constructive part of it. Students learn to argue while experiencing emotions linked to confrontation, error and the pressure to respond in real time. This is precisely what is missing in interaction with an artificial system, which produces arguments without undergoing any vulnerability. Goldberg and Schwarz specify that emotions should not be suppressed for reasoning to work, but oriented. When discussion is designed with pedagogical awareness, emotional reactions can become a lever for understanding more deeply the positions at stake, including one's own. This review therefore proposes to read the affective dimension of Debate not as a secondary component of argumentative practice, but as one of the conditions through which students learn to inhabit disagreement responsibly.

In the Italian context, Cinganotto shows how Debate is already understood in teaching as a methodology that integrates linguistic, social and socio-emotional dimensions. Mosa and Panzavolta argue that, with adequate design and differentiated roles, Debate can become an inclusive device in which even those who are not accustomed to public speaking find a structured and safe context in which to practise it.

In light of these contributions, Debate should not be understood as automatically inclusive or affectively formative; rather, it becomes so only when it is intentionally designed to support listening, gradual exposure, role differentiation and reflection on the emotional dynamics of discussion.

The third result can therefore be formulated as follows: Debate may be understood as a practice of affective digital literacy because it educates students to what AI can simulate but cannot live. An artificial system can construct relevant arguments, but it does not answer for those arguments before a real interlocutor who is waiting. Debate, by contrast, requires students to remain within the relationship, to listen genuinely,

to wait for the right moment to respond and to control the impulse to react before understanding. In theoretical-interpretive terms, this review proposes to read Debate as a pedagogical counterweight to the automatic production of language: it slows speech down, exposes it to responsibility and prevents language from becoming a pure flow separated from the experience of the person who produces it.

6. Discussion and pedagogical implications

The three results suggest a precise position: schools must not choose between using AI and resisting it, but must learn to distinguish what AI can do from what only human relational experience can produce. Generative AI offers useful resources for teaching, helping to construct examples, propose counterarguments and make otherwise inaccessible points of view available. The problem begins when it becomes a shortcut that replaces the process of thinking, searching for the right words or responding responsibly.

From this perspective, Debate should be understood as a complementary and critical device. It is complementary because it can integrate AI into activities of analysis, comparison and metacognition. It is critical because it restores to speech the dimensions that artificial communication tends to render invisible.

A teaching sequence can make this articulation concrete. In a first phase, students ask a generative system to produce arguments for and against a motion, not to use them, but to study them. In a second phase, the richest from a formative point of view, they analyse the texts produced, looking for persuasive strategies, emotional tones, implicit assumptions and neglected positions. In a third phase, they build their own Debate speech, deciding what to accept, reject or rework. In a fourth phase, they debate in person. Finally, metadialogue follows through targeted questions: what changed between reading an artificially generated argument and defending a position before a real interlocutor? At what point did the opponent's objection force one to think differently? The aim is to teach students not to confuse the automatic production of language with responsibility for speech.

Several pedagogical implications derive from this approach. The first concerns the redefinition of digital literacy. Verifying sources and using tools correctly are necessary competences, but not sufficient. It is also necessary to develop the capacity to recognise the emotional and persuasive effects of artificially generated communication, because an

artificial response may appear empathic and authoritative precisely when it requires critical analysis that goes beyond factual verification.

The second implication concerns the design of Debate as an educational practice, not as a competition. Debate conceived as an oratorical competition risk reinforcing the rhetoric of victory, whereas Debate designed pedagogically becomes a space for listening, metadialogue and conflict regulation, with gradual exposure to public speaking, clear criteria and attention to the emotional dynamics of the class.

The third implication concerns the role of the teacher, who, beyond assessing students' oral performance, becomes the director of the communicative environment, called upon to build psychological safety, guide dissent without allowing it to degenerate and make visible the emotions that run through discussion. The teacher's competence expands towards a mediation that is both affective and digital, integrating disciplinary competences with a new and urgent dimension.

The fourth implication concerns democratic citizenship. In an age in which AI can produce persuasive messages on a massive scale, personalised and difficult to distinguish from human ones, education for public speech and for listening to dissent becomes a civic competence in the most precise sense of the term. Debate is thus confirmed as one of the few educational environments in which this competence is practised systematically, under regulated conditions and with real interlocutors.

The fifth implication concerns inclusion. With differentiated roles, cooperative preparatory phases and alternating moments of writing and orality, Debate can become a space in which even those not accustomed to public speaking find an entry point into responsible expression.

The assessment rubric must also reflect this expansion. Alongside the argumentative dimension (clarity of thesis, quality of evidence, capacity for rebuttal) and the communicative dimension (use of voice, time management), it is necessary to include the affective-relational dimension (listening, emotional regulation, recognition of the interlocutor as a person distinct from the argument) and the metacognitive dimension: the capacity to reflect on one's own process, recognise one's biases and distinguish an argument personally elaborated from one uncritically taken from an external source. This last dimension is perhaps the most urgent: in the age of generative AI, metacognition about the argumentative process becomes inseparable from awareness of the tools used.

7. Limits of the review

Every literature review carries inevitable limits. In the case of this contribution, they are linked to the very nature of the territory explored: a field in which scientific research is expanding rapidly and remains, to a large extent, not yet consolidated.

The first limit concerns precisely the maturity of the field: studies on generative AI as a communicative agent are very recent, some published as preprints, and some of the contributions included may be updated by later studies. The conclusions describe the current state of the evidence, not a definitive framework. The second limit is the interdisciplinary nature of the synthesis: the relationship between generative AI and Debate touches cognitive science, pedagogy, linguistics and the ethics of technology, and the review cannot claim to have exhausted any of the individual fields involved. It is a thematic and interpretive synthesis, not a meta-analysis with uniform criteria. The third limit concerns the heterogeneity of the contributions included: alongside peer-reviewed empirical studies, there are policy reports and didactic-operational materials, necessary to contextualise the proposal within the regulatory framework. Finally, the fourth limit is theoretical: the category of affective digital literacy is not yet stabilised in the scientific literature, nor do validated instruments exist to measure it. This review is essentially propositional: it aims to identify and relate evidence from different fields in order to build the foundations of a category that future research will have to validate empirically.

8. Conclusion

What changes, for schools, when an algorithm learns to construct convincing arguments, respond in an empathic tone and adapt to the profile of the interlocutor? The assumption on which many of their most consolidated practices rested changes. If argumentative competence could be measured by the quality of a written text, today that quality can be produced artificially. If empathy expressed in a response was the sign of an authentic relationship, today it may be the result of a linguistic pattern learned from billions of words. If persuading someone required presence and the capacity to read the interlocutor, today it can occur through personalised messages generated on a massive scale.

The question that follows is precise: what does educational experience produce that no artificial system can generate? The answer this review has tried to build is that the formative processes irreproducible by AI are those that require presence, vulnerability, responsibility and relationship. The risks identified in the literature, from cognitive delegation to the standardization of language, are not inevitable; they become probable when AI is introduced without mediation and without spaces in which students are called upon to experience the authentic work of thinking in the first person. A school that designs well can transform the presence of AI into a formative opportunity, working precisely from the awareness of those risks.

From this perspective, Debate is one of the few educational environments in which conflict is regulated, listening is structurally necessary, speech is exposed to the verification of the other, and the emotions of dissent become an explicit object of formation. It restores to communication what automatic generation renders invisible: the time needed to find the right words, the body that feels the pressure of the moment, the voice that carries a history, the responsibility of answering in the first person.

The final thesis of this contribution can therefore be formulated in both descriptive and normative terms. The more algorithms learn to speak, the more schools are called to teach students to inhabit speech, not merely to produce it. Future research will need to empirically verify whether, and under what conditions, Debate pathways integrated with activities for the critical analysis of generative AI can effectively develop affective digital literacy skills. These dimensions cannot be transmitted through explanation alone; they are built through repeated, reflected and relationally situated experience. This is precisely the kind of experience that Debate, when intentionally designed as a pedagogical practice, can offer.

References

Asterhan, C. S. C., Howe, C., Lefstein, A., Matusov, E., & Reznitskaya, A. (2020). Controversies and consensus in research on dialogic teaching and learning. *Dialogic Pedagogy, 8*, S1–S16.

Biagini, G. (2025). Towards an AI-literate future: A systematic literature review exploring education, ethics, and applications. *International Journal of*

Artificial Intelligence in Education. Advance online publication. https://doi.org/xxxxx

Cinganotto, L. (2021). The Handy Little Guide to Debate: How to foster debating skills in the EFL class. Pearson.

Creely, E., & Carabott, K. (2025). Teaching and learning with AI: An integrated AI-oriented pedagogical model. The Australian Educational Researcher, 52(6), 4633–4654.

Galli, F. (2025). Intelligenza artificiale, didattica e regolazione: opportunità, rischi e prospettive normative. *Rivista italiana di diritto pubblico comunitario, 1*, xx–xx.

Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. Societies, 15(1), 6.

Goldberg, T., & Schwarz, B. B. (2016). Harnessing emotions to deliberative argumentation in classroom discussions on historical issues in multi-cultural contexts. Frontline Learning Research, 4(4), 7–19.

Guamanga, M. H., Saiz, C., Rivas, S. F., & Morales Bueno, P. (2025). Critical thinking and metacognition: pathways to empathy and psychological well-being. Journal of Intelligence, 13(3), 34.

Habermas, J. (1986). *Teoria dell'agire comunicativo*. il Mulino. (Opera originale pubblicata nel 1981)

Hackenburg, K., Tappin, B. M., Röttger, P., Hale, S., Bright, J., & Margetts, H. (2024). *Evidence of a log scaling law for political persuasion with large language models*. arXiv. https://arxiv.org/abs/2406.14508

Lee, H. P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM.

Lee, Y. K., Suh, J., Zhan, H., Li, J. J., & Ong, D. C. (2024). Large language models produce responses perceived to be empathic. In *Proceedings of the 2024 12th International Conference on Affective Computing and Intelligent Interaction (ACII)* (pp. 63–71). IEEE.

Ministero dell'Istruzione e del Merito. (2025). *Linee guida per l'introduzione dell'intelligenza artificiale nelle istituzioni scolastiche*. https://www.mim.gov.it/...

Moon, K., Green, A. E., & Kushlev, K. (2025). Homogenizing Effect of Large Language Models on Creative Diversity. Computers in Human Behavior: Artificial Humans, 100207.

Mosa, E., & Panzavolta, S. (2023). Diventare competenti con il Debate. RicercAzione, 15(2 bis), 143–161.

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and validation of the AI Literacy Questionnaire: The affective, behavioural, cognitive and ethical approach. British Journal of Educational Technology, 55(3), 1082–1104.

Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning. *Contemporary Educational Psychology, 33*(3), 345–359. https://doi.org/xxxxx

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41*(10), 994–1020. https://doi.org/10.1002/tea.20035

Perelman, C., & Olbrechts-Tyteca, L. (1958). *Traité de l'argumentation: La nouvelle rhétorique*. Presses Universitaires de France.

Rapanta, C. (2019). Argumentation as Critically Oriented Pedagogical Dialogue. Informal Logic, 39(1), 1–31.

Rapanta, C., & Macagno, F. (2025). From dialogue to metadialogue in classroom discussions. *Learning in Context, 1*(1–2), 100002. https://doi.org/10.1016/j.lecon.2025.100002

Rogiers, A., Noels, S., Buyl, M., & De Bie, T. (2024). *Persuasion with large language models: A survey*. arXiv. https://arxiv.org/abs/2411.06837

Salvi, F., Horta Ribeiro, M., Gallotti, R., & West, R. (2025). On the conversational persuasiveness of GPT-4. Nature Human Behaviour, 9, 1645–1653.

Savoldi, B., Attanasio, G., Gorodetskaya, O., Marchiori Manerba, M., Bassignana, E., Casola, S., Negri, M., Caselli, T., Bentivogli, L., Ramponi, A., Muti, A., Balbo, N., & Nozza, D. (2025). *Generative AI practices, literacy, and*

divides: An empirical analysis in the Italian context. arXiv. https://arxiv.org/abs/2512.03671

Sorin, V., Brin, D., Barash, Y., Konen, E., Charney, A., Nadkarni, G., & Klang, E. (2024). Large Language Models and Empathy: Systematic Review. Journal of Medical Internet Research, 26, e52597.

Sourati, Z., Ziabari, A. S., & Dehghani, M. (2026). The homogenizing effect of large language models on human expression and thought. *Trends in Cognitive Sciences, volume*(issue), pages. https://doi.org/xxxxx

Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.

Tortuga, & Yellow Tech. (2025). *Generazione AI: La nuova sfida della scuola*. [Report]

Turkle, S. (2012). Alone Together: Why We Expect More from Technology and Less from Each Other. Basic Books.

UNESCO. (2025). *AI in the classroom: Empowering teachers for the future of learning*. UNESCO.

Welivita, A., & Pu, P. (2024). *Is ChatGPT more empathetic than humans?* arXiv. https://arxiv.org/abs/2403.05572

Wilkie, K., & Ayalon, M. (2023). Learning to argue while arguing to learn. EURASIA Journal of Mathematics, Science and Technology Education, 19(8), em2312.

Yang, R., et al. (2025). Enabling inclusive systematic reviews. Journal of the American Medical Informatics Association, 32(11), 1718–1725.

Zuboff, S. (2023). The age of surveillance capitalism. In W. Longhofer & D. Winchester (Eds.), *Social theory re-wired: New connections to classical and contemporary perspectives* (pp. 203–213). Routledge.