CAN CARE BE SIMULATED? TEACHERS' JUDGMENTS OF EDUCATIONAL MESSAGES IN THE AGE OF GENERATIVE AI

Riccardo Sebastiani
Link Campus University
r.sebastiani@unilink.it

Riccardo Mancini
Link Campus University
r.mancini@unilink.it

Sara Pellegrini
Link Campus University
s.pellegrini@unilink.it

Double Blind Peer Review

Citation
Sebastiani, R., Mancini, R., & Pellegrini, S. (2026). Can care be simulated? Teachers' judgments of educational messages in the age of generative AI. *Italian Journal of Health Education, Sport and Inclusive Didactics, 10*(3).

Doi: https://doi.org/10.32043/gsd.v10i3.1734

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
This study explores how teachers and pre-service teachers interpret and evaluate educational messages when it is unclear whether they are written by a human or generated by artificial intelligence. Through a survey involving 807 participants in teacher qualification pathways, three scenarios and their corresponding author attributions were examined. The findings show that messages can be perceived as useful and plausible, yet recognition as expressions of educational care shifts when the origin of the text is considered. For the same content, assuming a human author tends to raise relational evaluations, while attributing the message to AI lowers them. As messages approach more sensitive situations, linguistic quality alone is insufficient: responsibility, boundaries, and a recognizable educational presence become central. These findings expose the tension between technophobic resistance and uncritical AI acceptance, placing teachers between the 'Frankenstein Complex' and the D-Generation.

KEYWORDS
Generative AI; teacher education; educational care.

Received 2026-06-12
Accepted 2026-10-22
Published 2026-10-30

1.Introduction

The introduction of generative artificial intelligence into education has opened an unstable space in which established practices and emerging possibilities intersect. This shift does not simply involve new tools, but affects how knowledge is produced, relationships are constructed, and value is assigned to educational language (Prilop et al., 2025). Recent studies show that teachers recognize the potential of generative AI to support instructional design, personalize materials, and accompany learning processes, while maintaining a cautious stance shaped by ambivalence and tension (Sebastiani, 2025).

Teachers' responses to AI appear to be shaped neither by outright rejection nor by full acceptance, but by an ongoing negotiation grounded in everyday teaching practice. They acknowledge its operational usefulness, while continuing to regard some dimensions of educational work as non-substitutable, especially those linked to students' personal development and the quality of the educational relationship (Chan & Tsi, 2024). The focus thus shifts from what AI can perform to what education still recognizes as irreducibly human.

Within this transition, language assumes a particularly sensitive role. Generative models can now produce texts that imitate empathic registers, affective nuances, and forms of encouragement, often being perceived as useful, relevant, and sometimes reassuring. Recent studies indicate that AI-generated content may be evaluated as educationally effective and comparable to teacher-produced content in terms of comprehensibility and perceived usefulness (Lee & Song, 2024; Netland et al., 2025). Yet the literature warns against conflating communicative effectiveness with educational quality, stressing that the relational dimension cannot be reduced to linguistic form alone (Dong, 2026).

At this point, the theoretical framework becomes more complex. Educational care cannot be reduced to message correctness or immediate effectiveness; it emerges as a situated practice involving responsibility, interpretive judgment, and attention to contexts and fragile relational signals. The adoption of AI in education makes this distinction particularly visible: while technological integration may expand instructional possibilities, it also requires a conscious redefinition of professional boundaries and teachers' responsibilities (Di Liso et al., 2025). The issue, then, is not only whether AI "works", but how far it can enter educational practices involving presence, judgment, and responsibility.

The perception of AI is not shaped linearly by demographic variables or professional experience, but by the interplay between trust, distrust, and

familiarity of use (Akanzire, 2025). Teachers' evaluations cannot therefore be explained only through external factors; they emerge from situated judgments formed in relation to texts, contexts, and everyday educational decisions.

An underexplored area of inquiry concerns not only the objective quality of AI-generated texts, but the extent to which readers perceive them as relationally plausible and connected to educational care. Evidence suggests that teachers may evaluate such content positively while maintaining distance from its relational legitimacy (Velez et al., 2025). What remains unclear is how these evaluations change when the text's origin is uncertain or explicitly attributed to either a human or artificial author.

The present study is situated within this gap. It does not take a normative position on the use of AI, nor does it assess whether such use is appropriate. Rather, it examines what happens when a relationally significant message is read without knowing its author, and how its evaluation changes when the same text is attributed to either a teacher or an artificial system. In education, care is not defined by textual quality alone, but by the relationship between words, authorship, and responsibility. On this basis, the study explores how perceptions of educational care vary according to human or artificial attribution.

2.Methodology

Consistent with the research question, the study adopts an exploratory, quantitatively driven mixed-methods design with an integrated qualitative component, aimed at examining how judgments of educational messages vary according to their attribution as human, artificial, or uncertain. Since the questionnaire was an original instrument developed for this study, it was used exploratorily; the data are not treated as evidence of full validation, but as a preliminary basis for assessing theoretical plausibility, empirical readability, and internal coherence. The aim is to understand how teachers and participants in teacher qualification pathways evaluate the relational quality of educationally dense messages when authorship enters the judgment process. The study therefore addresses not only the acceptability of AI in education, but also perceived authenticity, educational responsibility, boundary maintenance, acceptance, and the recognition of vulnerability.

The mixed-methods approach integrates Likert-scale quantitative judgments with a qualitative component designed to identify the criteria through which participants define and delimit "care" in education (Likert, 1932). Data were collected through an anonymous and voluntary online questionnaire administered between January 15 and April 1, 2026, to participants enrolled in

teacher qualification pathways at Link Campus University. Completion was preceded by an information sheet outlining the study aims, data processing procedures, and optional nature of participation. The non-probabilistic convenience sample allows the study to capture orientations, interpretive tensions, and descriptive regularities within the investigated context, but does not support statistical generalization to the wider teaching population.

A total of 807 course participants took part in the study, although the number of valid cases varies across questionnaire sections due to missing responses or partial completions, a common feature of web surveys that should be methodologically acknowledged (Eysenbach, 2004). The sample is predominantly female, concentrated in intermediate age groups, and mainly composed of upper secondary school teachers, most of whom were enrolled in the 30-ECTS teacher qualification pathway pursuant to Article 13. A substantial proportion also reported frequent, often almost daily, use of AI for study, practicum activities, or instructional preparation. These background data situate the judgments collected within a population already familiar with both educational practice and the growing presence of AI in training and educational contexts.

The questionnaire was organized into integrated sections, aligned with the study's theoretical framework and exploratory aim. The scenarios and stimulus messages were developed ad hoc by the authors, drawing on recent literature on generative AI, educational relationships, and teachers' judgment, in order to construct plausible situations differentiated by relational intensity (Chan & Tsi, 2024; Luo, 2025; Moorhouse & Kohnke, 2024; Prilop et al., 2025; Tripathi et al., 2025). The first section collected demographic and professional data, the second examined the frequency and domains of generative AI use, and the third, constituting the core of the instrument, presented three educational vignettes with response messages designed to appear credible within school practice (Aguinis & Bradley, 2014; Atzmüller & Steiner, 2010). The situations concerned, respectively, assessment anxiety, sensory overload in a student with special educational needs, and an evening request for help from a student experiencing loneliness and family difficulties.

In the project coding, the messages for Scenarios 1 and 3 were AI-generated, while the message for Scenario 2 was human-authored. This information was not disclosed to participants during questionnaire completion and was used only during analysis. After reading each message, participants rated six dimensions: perceived acceptance, authenticity of tone, educational responsibility, usability as an initial form of support, risk of confusion or boundary shifting in the educational relationship, and educational validity even

under the hypothesis of artificial generation. For each scenario, they were also asked to attribute the message to a teacher/human being, an artificial intelligence system, or to indicate inability to decide.

For the analysis, a synthetic indicator of the message's relational quality was constructed as the mean of acceptance, authenticity, educational responsibility, initial support, and boundary maintenance, after reversing the negatively worded item on relational boundary crossing. This indicator does not exhaust the complexity of educational care, but enables a synthetic and comparative reading of evaluations across the three scenarios.

Closed-ended dimensions were measured using a five-point Likert scale, widely adopted in attitude studies because it captures both the direction and perceived intensity of judgment (Likert, 1932). Quantitative data were analyzed through descriptive statistics, including absolute frequencies, percentages, and response distributions by item and scenario. This choice reflects the exploratory nature of the study: the aim was not to develop a causal model or test strong inferential hypotheses, but to report emerging trends, concentrations of judgment, and discrepancies between educational appreciation of the message and perceived human or artificial authorship.

In an exploratory study, methodological robustness depends less on sophisticated statistical procedures than on the coherence between the research question, data structure, and interpretive caution. The analysis was therefore aligned with the dataset and the exploratory function of the design. Inferential analyses were used exploratorily, not to support generalizable conclusions, but to examine the internal consistency of descriptive patterns and the distribution of differences across scenarios and attribution groups.

Open-ended responses were analyzed through reflexive thematic analysis (Braun & Clarke, 2006), focusing on the textual answers provided to the three final open-ended questions. The corpus was examined through repeated readings, initial open coding, and progressive clustering of codes into thematic cores, defined according to recurrence and relevance to the research question. Coding proceeded iteratively through rereading, aggregation, and refinement until the most recurrent and theoretically relevant interpretive categories were stabilized.

This methodological choice is particularly appropriate for semantically dense and morally exposed constructs such as care, empathy, listening, responsibility, judgment, and the educational relationship. It avoids reducing responses to predefined labels and allows the grammars through which participants distinguish delegable from non-delegable teaching actions to emerge. The

qualitative component was therefore not used as mere support for the quantitative data, but as an autonomous interpretive space for clarifying how participants define educational care and delimit what may or may not be delegated to an artificial system.

The methodological framework was guided by constructive caution, ensuring coherence among the theoretical construct, the instrument structure, and the interpretive use of data. The procedure followed standards that require this connection to be made explicit, especially when investigating educational and psychological dimensions that are not directly observable (AERA, APA, & NCME, 2014). The survey was also reported according to transparency criteria for online research, specifying voluntary participation, variability in response rates, and limits related to sample self-selection (Eysenbach, 2004).

The central methodological issue is to preserve the exploratory nature of the study without overextending its inferential scope, while maintaining the rigor needed to open a credible line of research on the relationship between linguistic simulation, recognition of care, and educational responsibility.

3.Results

The results outline a participant profile marked by widespread familiarity with generative tools and a strong connection to school practice. Women were predominant in the sample (74.96%), with higher concentrations in the 35–44 and 45–54 age groups. Most respondents worked in upper secondary schools and were enrolled in the 30-ECTS teacher qualification pathway pursuant to Article 13. More than half held annual substitute contracts or appointments until the end of teaching activities, indicating a relatively stable position within the school context.

Almost two out of three participants reported near-daily educational contact with students, situating their judgments within a continuous rather than occasional school relationship. AI use was also far from marginal: a substantial proportion reported using it at least weekly, mainly for study and lesson preparation. While this does not explain the response pattern in itself, it weakens the idea that participants' judgments stemmed from theoretical distance or preconceived technological rejection. Across the three scenarios, the results do not show a clear opposition between acceptance and rejection. The messages were often perceived as relevant and relationally plausible, although agreement decreased when they entered more delicate areas, where verbal support alone appeared less sufficient.

In the first scenario, concerning a student overwhelmed by anxiety during an assessment, the values remain highest on the core dimensions. Acceptance

reaches 85.32% in the two highest agreement categories, educational responsibility 83.08%, and willingness to use that response as an initial form of support 77.88%. Authenticity also remains at a substantial level, with 65.01% of responses falling within the two upper categories. When responses are expressed on a scale from 1 to 5, the results appear rather high: 4.19 for acceptance, 4.09 for educational responsibility, 3.93 for initial support, and 3.58 for authenticity (Figure 1).

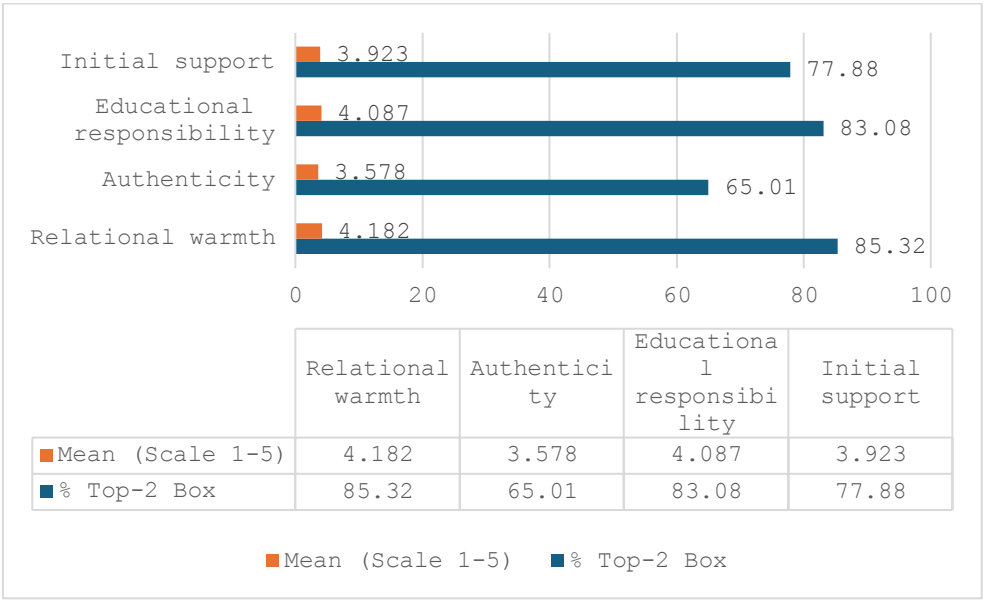


Figure 1: Comparative pattern of the three scenarios across the four common dimensions

In terms of authorial attribution, 42.11% of participants imagined an artificial intelligence system, 31.58% attributed the text to a teacher, and 26.32% stated that they were unable to decide. A divergence thus emerges between the evaluation of the message and the attribution of its author (Figure 2).

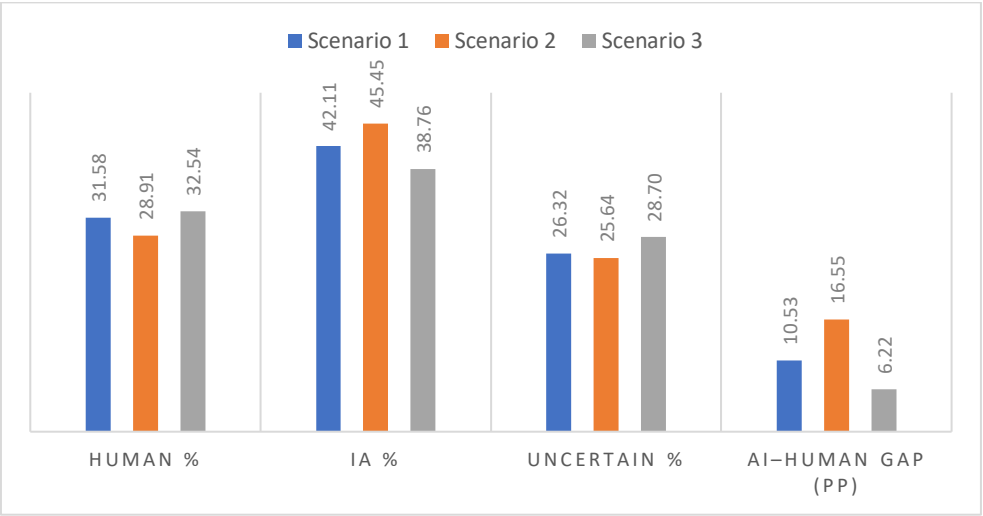


Figure 2: Author attribution across the three scenarios: human teacher, artificial intelligence system, and uncertainty.

In the second scenario, which in the project coding corresponds to the human-authored text, the profile remains solid but takes on a slightly different configuration, since the level of agreement, while still high, no longer shows the same compactness as in the first scenario. The values of 70.88% for acceptance, 69.01% for educational responsibility, and 63.97% for the possibility of using the response as an initial form of support continue to outline an overall positive picture. What emerges is not a real decline, but rather a less compact distribution of responses compared with the previous scenario, suggesting greater internal differentiation in participants' judgments. The weighted means follow the same pattern and confirm a good degree of consistency: 3.81 for acceptance, 3.78 for educational responsibility, 3.58 for initial support, and 3.35 for authenticity.

These values remain positive, but they are less compact, especially with regard to perceived authenticity. One element remains strongly anchored in the educational dimension: the item concerning the possible crossing of educational boundaries is rejected by slightly more than half of the respondents. In its reversed form, the synthetic indicator of the message's relational quality reaches a mean of 3.45, indicating that the text is perceived as more regulated and less intrusive. This scenario also shows the highest proportion of attribution of the message to AI: 45.45% of participants attributed the message to AI, 28.91% to a human teacher, while 25.64% remained in a zone of uncertainty. The most relevant finding is that the human-authored text was, at the perceptual level, the one least frequently recognized

as such. A gap thus emerges between the actual origin of the message and participants' perceived attribution.

In the third scenario, centred on an evening request for help related to loneliness and family difficulties, the evaluative profile changes more markedly. Judgments remain positive overall, but they are more widely distributed. The highest agreement categories decrease to 58.58% for acceptance, 58.24% for educational responsibility, and 48.59% for initial support; authenticity reaches 48.14%. The means are 3.47 for acceptance, 3.52 for educational responsibility, 3.15 for initial support, and 3.21 for authenticity. The scores therefore remain at moderate levels, but with less compactness than in the previous scenarios. In this scenario, the issue of boundary maintenance emerges more clearly: 40.63% of participants fall within the highest agreement categories with the idea that the text may blur the boundaries of the educational relationship, while 35.99% distance themselves from this interpretation (Figure 3).

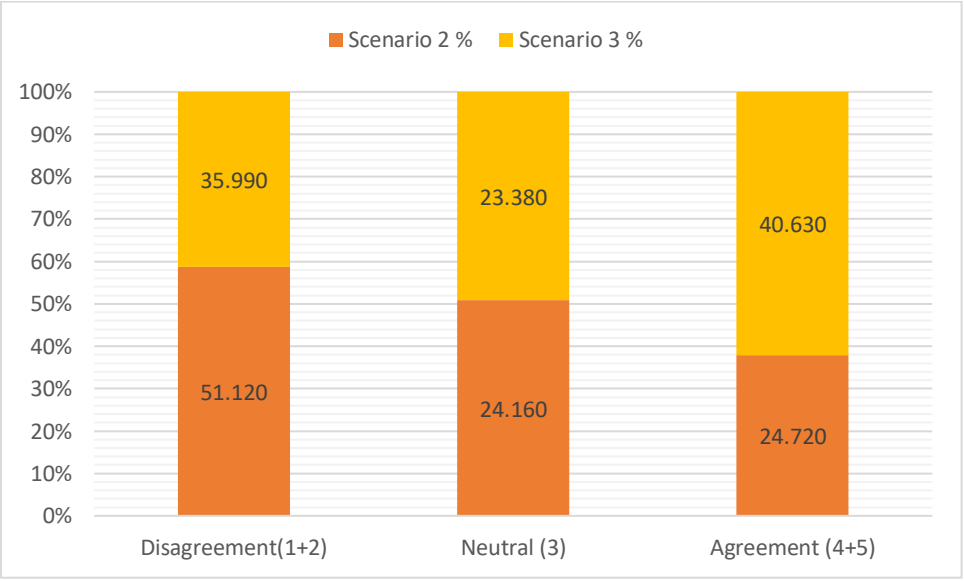


Figure 3. Perception of the boundaries of the educational relationship in Scenarios 2 and 3.

The educational validity of the text, even under the hypothesis of its artificial origin, reaches a mean of 3.16, a non-negligible value but lower than the other dimensions considered. Author attribution also does not reveal a clear-cut division: 38.76% attributed the message to AI, 32.54% to a teacher, while 28.70% remained uncertain.

When the analysis moves to a more systematic reading, the overall pattern is confirmed. The Friedman test indicates significant differences among the three scenarios across all the dimensions considered: acceptance ($\chi^2(2) = 167.45$, $p < .001$), authenticity ($\chi^2(2) = 40.37$, $p < .001$), educational responsibility ($\chi^2(2) = 113.75$, $p < .001$), initial support ($\chi^2(2) = 134.30$, $p < .001$), boundary maintenance ($\chi^2(2) = 92.97$, $p < .001$), and educational validity under the hypothesis of artificial origin ($\chi^2(2) = 71.34$, $p < .001$).

When considering the synthetic indicator of the message's relational quality, constructed from acceptance, authenticity, educational responsibility, initial support, and boundary maintenance, after reversing the item concerning the risk of relational boundary crossing, the difference among the three texts emerges clearly. The statistical analysis confirms a differentiation that was already visible at the descriptive level.

Kendall's coefficients are small (W ranging from .013 to .053), indicating effects of limited magnitude but consistent in their direction.

The three scenarios are not positioned at the same level: the first stands out quite clearly, the second follows closely behind, while the third remains lower. This difference is also confirmed statistically (Figure 4).

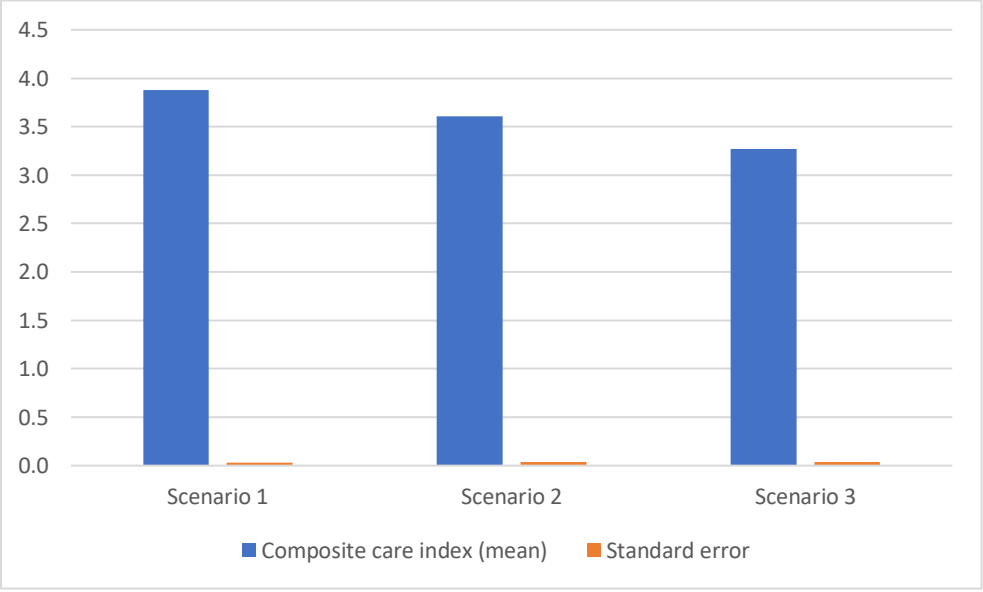


Figure 4: Synthetic indicator of the message's relational quality across the three scenarios

The post-hoc comparisons, conducted as paired analyses, confirm the same direction. The synthetic indicator of the message's relational quality is higher in the first scenario than in the second (Wilcoxon $V = 19357$, $p < .001$), higher in the first than in the third ($V = 12441.5$, $p < .001$), and likewise higher in the second than in the third ($V = 23525$, $p < .001$).

In terms of internal consistency, Cronbach's α for the synthetic indicator of the message's relational quality reaches .759 in the first scenario, .798 in the second, and .781 in the third. When the item concerning educational validity under the hypothesis of artificial origin is also included, the values rise to .801, .844, and .841, respectively. These coefficients, without suggesting perfect consistency, indicate a more than adequate level of reliability for an ad hoc exploratory instrument.

Perhaps the most revealing shift emerges, however, when the evaluation of the text is examined in relation to author attribution. The Kruskal–Wallis tests show, across all three scenarios, that the synthetic indicator of the message's relational quality varies significantly according to whether the message is attributed to a human teacher, to an AI system, or remains within a zone of uncertainty: Scenario 1, $H(2) = 47.49$, $p < .001$; Scenario 2, $H(2) = 126.81$, $p < .001$; Scenario 3, $H(2) = 110.99$, $p < .001$. The group means show a consistent gradation. In the first scenario, the mean value of the synthetic indicator of the message's relational quality is 4.13 among those who attribute the text to a human teacher, 3.95 among those who remain uncertain, and 3.63 among those who attribute it to AI. In the second scenario, the gap widens: 4.17 for human attribution, 3.64 for uncertainty, and 3.22 for attribution to AI. In the third scenario, the same pattern is confirmed: 3.80 for human attribution, 3.23 for uncertainty, and 2.87 for attribution to AI. The point here appears rather clear: for the same text read, imagining a human author systematically corresponds to a higher evaluation of its relational quality. In this sense, the Turing test does not concern only whether a message "passes" as human or not; it concerns the way in which that image of the author enters into the judgment and shifts it.

Spearman's correlations among the three synthetic indicators show a moderate continuity in the way messages are evaluated: Scenario 1 and Scenario 2 show $\rho = .454$ ($p < .001$), Scenario 1 and Scenario 3 show $\rho = .245$ ($p < .001$), while Scenario 2 and Scenario 3 show $\rho = .379$ ($p < .001$). These associations are not strong, but they are consistent, suggesting a degree of individual stability in the evaluation criteria. A similar pattern is also observed for the item concerning the educational validity of the message under the hypothesis of artificial origin:

ρ = .472 between Scenarios 1 and 2, ρ = .307 between Scenarios 1 and 3, and ρ = .303 between Scenarios 2 and 3, all with $p < .001$.

By contrast, general orientations toward AI are only weakly associated with the synthetic indicators of the message's relational quality, and the correlations with the mean of the three indicators across the scenarios remain low and non-significant. This is, in itself, an interesting finding. It seems to suggest that judgment is not determined in advance by an ideological attitude toward technology, but rather takes shape in the concreteness of the text and of the situation.

This impression is also consistent with the exploratory linear regression model in which the mean synthetic indicator of relational quality across the three scenarios was entered as the dependent variable, while frequency of AI use, intensity of educational contact, age, teaching experience, general attitude toward AI, and gender were included as predictors. The model shows an overall limited explanatory capacity and does not reach overall statistical significance ($R^2 = .023$, $F(6,357) = 1.37$, $p = .226$). The results are therefore treated as weak and preliminary indications, insufficient to support robust inferences regarding the weight of the individual predictors.

The final section of the questionnaire, devoted to orientations toward AI in education, reveals a cautious but not rigid attitude. Overall, 45.88% of respondents consider it a risk to be contained, 48.03% fear a possible weakening of the educational relationship, 42.88% would prefer to avoid it even when it appears useful because of the risk of dependency, and 49.31% believe that simulated empathy is not compatible with educational care. In this section as well, intermediate positions remain broad. What emerges is not an absolute closure, but rather a cautious, selective stance that is still in the process of stabilization. This stance is, in many respects, similar to that observed in other recent studies on the practices and forms of trust and distrust that accompany the entry of generative AI into educational contexts (Prilop et al., 2025) (Figure 5).

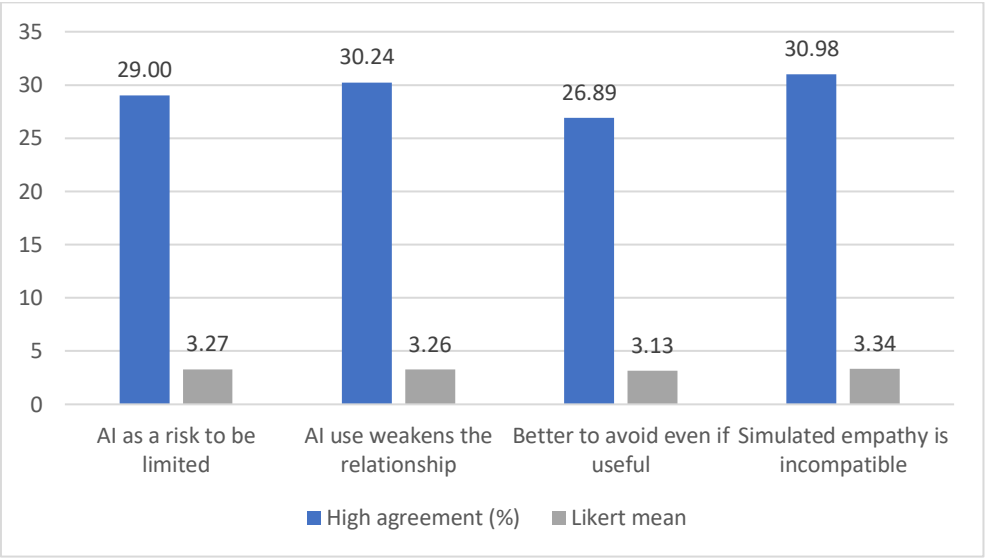


Figure 5. General attitudes toward AI

The qualitative material moves in the same direction and helps to clarify it. When participants were asked which elements make a sense of care visible in a response, terms such as listening, empathy, authenticity, attention to context, personalization, and the ability not to leave the student alone recur frequently. When, instead, they were asked what a teacher should never delegate to an artificial system, the responses converge around recurring cores such as empathy, the educational relationship, assessment, judgment on the learning pathway, the emotional sphere, human presence, and moral responsibility.

Some responses are extremely concise ("Empathy", "Listening", "Assessment"), while others are more elaborate, but the semantic core remains substantially convergent. AI is recognized as potentially useful for suggesting, organizing, or easing certain functions, but not as fully legitimate in occupying the most exposed dimensions of teachers' work. In several passages, this idea is expressed very directly: "the responsibility for the educational relationship cannot be delegated", "only the teacher can grasp subtle signals", "educating is, first and foremost, a human relationship".

Taken together, the qualitative analysis confirms a coherent pattern: artificial messages may be perceived as useful, well constructed, and at times reassuring, yet their full relational legitimacy appears more limited. The linguistic form of the message is relevant, but not sufficient to replace dimensions that participants continue to regard as decisive: presence,

judgment, responsibility, a sense of limits, and the capacity to inhabit the educational relationship without reducing it to a correct formula.

4.Discussion

Within the limits of the sample and research design, the results provide a plausible answer to the research question. The messages are often judged appropriate and, in some cases, useful as an initial form of support; their evaluation, however, changes systematically when author attribution is considered.

For the same text, human authorship is associated with higher relational evaluations, whereas attribution to AI corresponds to lower evaluations. Participants thus appear to recognize generative systems as capable of imitating the linguistic form of support, without attributing to them, with the same ease, the responsible presence associated with educational care. This also explains the distribution across the three scenarios: the first receives the highest evaluations, the second remains consistent, while the third, more exposed at a personal and relational level, elicits less stable consensus. As messages enter more delicate relational areas, writing quality alone becomes less sufficient to sustain their recognition as care; human recognizability, responsibility, and the capacity to combine proximity with educational boundaries become more decisive.

The data are consistent with much of the recent literature. Studies show that teachers and teacher educators recognize the usefulness of generative AI for instructional design, material personalization, and learning support, while continuing to view specifically human dimensions of teaching, pedagogical judgment, relationships, and professional responsibility, as non-substitutable (Chan & Tsi, 2024; Prilop et al., 2025). Research on teachers' perceptions confirms this ambivalence: openness to operational uses coexists with concerns about relational quality, equity, professional boundaries, and insufficiently reflective use of generative tools (Akanzire et al., 2025).

The study adds a more fine-grained element: it does not merely confirm teachers' caution, but shows where this caution is activated. It emerges less in relation to AI in general than when a text approaches the other's vulnerability. This aligns with research distinguishing trust and distrust as non-symmetrical dimensions: teachers may recognize a tool's usefulness while hesitating to entrust it with sensitive aspects of educational action (Tripathi et al., 2025). This also helps explain why, in the adopted model, general orientations toward AI are only weakly associated with the care indices. From this perspective, the

paradox of the second scenario is particularly significant: the human-authored text is the one most frequently attributed to AI.

A pedagogically composed, attentive, and internally regulated form of writing may now appear linguistically compatible with what participants associate with artificial production. In the contemporary professional imaginary, highly controlled and "clean" writing can therefore be read as suspiciously artificial. This finding supports the need for an AI literacy that is not only technical, but also critical and interpretive: teachers must understand not only how to use these tools, but how their diffusion reshapes linguistic expectations, criteria of authenticity, and the recognition of pedagogical authority in educational discourse (Prilop et al., 2025).

The qualitative material reinforces this interpretation. Terms such as listening, empathy, presence, attention to context, and moral responsibility define an educational stance that separates what can be supported from what cannot be delegated. AI may help, facilitate, and produce linguistically plausible responses, but it is not perceived as fully suited to sustaining the thresholds of sensitivity required when educational judgment involves fragility, limits, suffering, or the need to accompany without intruding. This distinction between functional support and relational legitimacy is also confirmed by recent studies, especially when teaching is understood as extending beyond content delivery (Chan & Tsi, 2024).

The concept of "educational care" remains semantically dense and open to multiple interpretations. In this study, dimensions such as acceptance, empathy, reliability, and responsibility tend to converge, although future research should distinguish them more precisely. The instrument's good internal consistency is encouraging, but it does not resolve the issue; rather, it calls for further investigation through more refined measures and more differentiated samples.

Theoretically, care depends not only on how a message is constructed, but on the relationship between words, attributed authorship, and the educational weight they carry. A text may appear compatible with a human register, yet this is not sufficient: it must also be perceived as trustworthy and legitimate within an educational relationship (Sebastiani et al., 2025). This has direct implications for teacher education. Teachers should not only learn to use generative AI as a technical support tool, but also to distinguish what can be delegated from what cannot, recognize relational boundaries, and develop a situated and conscious use of these technologies, as recent literature has emphasized (Prilop et al., 2025).

At the institutional level, the results call for avoiding both uncritical enthusiasm and prejudicial rejection; at the professional level, they point to the need for an AI use that does not obscure the relational core of teachers' work. These findings must be read in light of several limitations: the convenience-based sample, the specific teacher education context, the exploratory use of an original instrument, the vignette-based nature of the scenarios, and the possible influence of stylistic features in the stimulus messages. The evidence should therefore be understood not as generalizable, but as a preliminary empirical basis for further research.

Several questions remain open. Future research should examine whether the same pattern emerges across different school levels, non-qualification contexts, established teachers, and students. Further studies could manipulate not only author attribution, but also textual style, degree of personalization, and the presence of explicit boundary markers. More ecologically realistic designs, including longitudinal interactions, dialogic simulations, multimodal scenarios, or settings incorporating the relational memory of the situation, would also be valuable. A further issue concerns whether increasing cultural familiarity with AI-generated texts will reshape the criteria through which words are perceived as authentic or as "too perfect" to be human.

Taken as a whole, the study suggests that the evaluation of relationally dense educational messages depends not only on verbal content, but also on attributed authorship and the educational responsibility associated with it. Generative AI can produce texts perceived as plausible and sometimes useful, but participants do not automatically recognize them as equivalent to a human educational presence.

5.Conclusion

The data highlight an aspect often overlooked in debates on AI in education: in school contexts, the evaluation of a message depends not only on its content or immediate effectiveness, but also on who the reader believes has written it. This is the study's central finding: texts judged appropriate, and sometimes reassuring, are not necessarily recognized as expressions of fully trustworthy educational care.

The most pronounced finding concerns the gap between perceived message quality and author attribution. For the same text, hypothesized human authorship increases relational evaluation, whereas attribution to AI lowers it. This pattern does not seem to reflect simple distrust of technology, but the fact that, within educational relationships, words are not read as neutral

statements. They are interpreted as situated gestures, linked to responsibility, boundaries, and possible presence.

This interpretation calls for caution. Generative AI can produce plausible and well-calibrated messages, often useful as an initial form of support. Yet when situations involve personal fragility, suffering, requests for help, or boundary management, linguistic competence alone is insufficient. Even when formally appropriate, the text does not always carry the relational weight that education associates with care.

The paradox of the second scenario deserves particular attention: the human-authored text is the one most frequently attributed to AI. This is not a marginal detail, but an indication that, in the current cultural climate, highly composed, controlled, and pedagogically regulated writing may appear suspiciously artificial. The finding therefore concerns not only the tools used, but also the criteria through which a word is recognized, or no longer recognized, as authentically human.

The scope of these results should not be overstated: the sample is circumscribed, the instrument original, and the scenarios constructed rather than based on real interactions. Yet the study should not be reduced to a merely preliminary test. Despite its limitations, it raises a question that remains central for pedagogical research: not only whether AI can produce correct or well-written responses, but when a word is perceived as educational, legitimate, and capable of conveying responsible presence within a relationship. Education never coincides fully with the form of a message; it also depends on what that form makes perceptible, a stance, an assumption of responsibility, a measured sense of boundaries, and the capacity to accompany without occupying the other's space. On this terrain, AI reveals both its strength and its limit, and pedagogical research is called to work with greater precision and fewer simplifications.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods, 17*(4), 351–371. https://doi.org/10.1177/1094428114547952

Akanzire, B. N., Acquah, S., Nyaaba, M., Densu, A., & Tannor, P. (2025). Generative AI in teacher education: Teacher educators' perception and preparedness. *Journal of Digital Educational Technology, 5*(1), ep2508. https://doi.org/10.30935/jdet/15887

Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence, 6*, 100215. https://doi.org/10.1016/j.caeai.2024.100215

Alqarni, T. M. (2026). Teachers' perceptions of generative AI in inclusive classrooms: Enhancing engagement for students with learning disabilities. *Universal Access in the Information Society, 25*, Article 22. https://doi.org/10.1007/s10209-025-01301-8

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 6*(3), 128–138. https://doi.org/10.1027/1614-2241/a000014

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

Chan, C. K. Y., & Tsi, L. H. Y. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation, 83*, 101395. https://doi.org/10.1016/j.stueduc.2024.101395

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

Di Lisio, G., Milá Roa, A., Halty, A., Berástegui, A., Couso Losada, A., & Pitillas, C. (2025). *Nurturing bonds that empower learning: A systematic review of the significance of teacher-student relationship in education*. *Frontiers in Education, 10*. https://doi.org/10.3389/feduc.2025.1522997

Dong, Y. (2026). *Generative AI technologies and educational outcomes: A comprehensive meta-analysis comparing traditional and AI-driven approaches*. *Humanities and Social Sciences Communications*. Advance online publication. https://doi.org/10.1057/s41599-026-06903-y

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Eysenbach, G. (2004). Improving the quality of web surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *Journal of Medical Internet Research, 6*(3), e34. https://doi.org/10.2196/jmir.6.3.e34

Field, A. (2024). *Discovering statistics using IBM SPSS Statistics* (6th ed.). SAGE.

Lee, S., & Song, K.-S. (2024). *Teachers' and students' perceptions of AI-generated concept explanations: Implications for integrating generative AI in computer science education*. *Computers and Education: Artificial Intelligence, 7*, 100283. https://doi.org/10.1016/j.caeai.2024.100283

Likert, R. (1932). *A technique for the measurement of attitudes*. *Archives of Psychology, 140*, 1–55.

Luo, J. (2025). How does GenAI affect trust in teacher-student relationships? Insights from students' assessment experiences. *Teaching in Higher Education, 30*(4), 991–1006. https://doi.org/10.1080/13562517.2024.2341005

Moorhouse, B. L., & Kohnke, L. (2024). The effects of generative AI on initial language teacher education: The perceptions of teacher educators. *System, 122*, 103290. https://doi.org/10.1016/j.system.2024.103290

Netland, T., von Dzengelevski, O., Tesch, K., & Kwasnitschka, D. (2025). *Comparing human-made and AI-generated teaching videos: An experimental study on learning effects*. *Computers & Education, 224*, 105164. https://doi.org/10.1016/j.compedu.2024.105164

Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education*. University of California Press.

Noddings, N. (2005). *The challenge to care in schools: An alternative approach to education* (2nd ed.). Teachers College Press.

Prilop, C. N., Mah, D.-K., Jacobsen, L. J., Hansen, R. R., Weber, K. E., & Hoya, F. (2025). Generative AI in teacher education: Educators' perceptions of transformative potentials and the triadic nature of AI literacy explored through AI-enhanced methods. *Computers and Education: Artificial Intelligence, 9*, 100471. https://doi.org/10.1016/j.caeai.2025.100471

Sebastiani, R. (2025). *Educazione alla sostenibilità e innovazione didattica: verso un apprendimento inclusivo*. RicercAzione, 17(1). https://doi.org/10.32076/RA17105

Sebastiani, R., Cuccaro, A., Gentilozzi, C., & Pellegrini, S. (2026). *Inclusion as embodiment: Towards an integrated paradigm between embodied education, inclusive pedagogy and artificial intelligence*. In A. Dipace, C. Limongelli, M. Marras, & S. M. Pagliara (Eds.), *Artificial intelligence with and for learning sciences: Past, present, and future horizons* (LNCS 16438, pp. 156–171). Springer. https://doi.org/10.1007/978-3-032-17604-2_15

Teo, T. (2011). Factors influencing teachers' intention to use technology: Model development and test. *Computers & Education, 57*(4), 2432–2440. https://doi.org/10.1016/j.compedu.2011.06.008

Tripathi, T., Sharma, S. R., Singh, V., Bhargava, P., & Raj, C. (2025). *Teaching and learning with AI: A qualitative study on K-12 teachers' use and engagement with artificial intelligence*. *Frontiers in Education, 10*, 1651217. https://doi.org/10.3389/feduc.2025.1651217

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

Velez, G., Xu, L. Z., & Fitzgerald, J. (2026). *From AI as a relational breaker to a relational broker: Comment on Bauer et al. (2025)*. *Educational Psychology Review, 38*, Article 15. https://doi.org/10.1007/s10648-025-10106-

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540