



OPEN ACCESS

EDITED BY

Xi Chen,
Yunnan University, China

REVIEWED BY

Ling Huang,
Yunnan University, China
Afendi Hamat,
National University of Malaysia, Malaysia

*CORRESPONDENCE

Alina Roiniță
✉ alina.roinita@unibuc.ro;
✉ alina.roinita@drd.unibuc.ro

RECEIVED 19 October 2025

REVISED 13 January 2026

ACCEPTED 19 January 2026

PUBLISHED 05 February 2026

CITATION

Roiniță A (2026) Human touch versus algorithm: reception of AI poetry among Romanian adolescents.
Front. Educ. 11:1728281.
doi: 10.3389/feduc.2026.1728281

COPYRIGHT

© 2026 Roiniță. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Human touch versus algorithm: reception of AI poetry among Romanian adolescents

Alina Roiniță*

Centre of Romanian Studies, University of Bucharest, Bucharest, Romania

This research paper presents the results of a case study on how Romanian adolescents perceive poetry written by Romanian authors versus poetry generated by an artificial intelligence (AI) model. The study mainly relied on quantitative methods (descriptive statistical analysis using SPSS) and interpretive discussion of trends emerging from the aggregated response data. It involved a control group, which was informed from the beginning about the research variables, including the fact that they would evaluate both human-written poems from various literary movements and AI-generated poems produced using prompts based on texts from the same poetry volumes, and an experimental group, which did not know the texts' origins. The experimental group was only informed about the origin of the texts at the end, at which point they were invited to re-evaluate the AI-generated poems. The aim of this study is to investigate how adolescents perceive and evaluate human-written poetry from different literary movements compared to AI-generated poetry. The analysis examines differences in emotional responses, perceived creativity and complexity of the poems, and how these perceptions are shaped by the author's identity (AI or human), as well as by the respondents' educational background and gender. The main findings reveal that adolescents' perception of poetry is strongly shaped by authorial labels, with human poems receiving more favourable evaluations when their origin is known. It is worth noting that, in the absence of such information, AI-generated poems were often rated higher than those written by human authors. Once the authorship was revealed, however, AI-generated texts tended to be penalized – especially by female respondents and those with a humanities background – possibly suggesting a stronger attachment to “traditional literature” that influences their reception. Emotional responses also vary with authorship: human poems tend to be more often associated with positive emotions like happiness, while AI-generated texts tend to trigger negative emotions, particularly when inspired by specific literary movements and authors.

KEYWORDS

AI-generated poetry, authorship effect, digital education, emotional response, high school adolescents, literary reception

1 Introduction

In recent years, natural language processing and understanding models have advanced fast, leading to the emergence of Large Language Models (LLMs), capable of generating text based on user input (named “prompts”) after being trained on extensive texts corpora. These LLMs form the backbone of different AI applications such as ChatGPT, Deepseek and Gemini, which can produce increasingly fluent texts adapted to both context and interlocutor. Built on artificial neural networks inspired by the structure of the human brain – particularly “transformer” architectures, AI models are playing an increasingly

prominent role in societal development (Watson and Romic, 2024). Referring specifically to ChatGPT, Watson and Romic argue that “ChatGPT not only performs tasks, but also mediates thought, action, and communication.” They further argue that these models may assume a central societal role, as they not only assist cognitive processes but also “functions as part of a dynamic semantic ecosystem, structurally coupling thought with action and society” (Watson and Romic, 2024). This perspective on the development of AI applications helps to explain why LLMs have become so influential in so many different areas of our contemporary society. Not only they process and generate information, but also shape the way humans think, communicate and create meaning – and why not, take decisions.

In recent years, AI models have been applied across numerous fields, including medicine, economics, education, not only to streamline tasks but also to support thinking, planning, collaboration and decision-making. However, the fast expansion of AI applications has triggered numerous ethical concerns, particularly regarding research integrity (Christou, 2023), data fabrication (Limongi, 2024), and challenges to traditional notions of authorship (Ebadi et al., 2025). These ethical concerns can be particularly pertinent in fields where creativity and authorship are fundamental, such as humanities.

The increasing use of AI in humanities and creative fields, including literature, raises important questions regarding authorship and aesthetic value nowadays. This is because literary creation depends on language, area in which AI is playing an important role. As Grace and Sahu (2024) highlight, there are multiple crossovers between AI and literature – from content generation and language enhancement to character creation, dialogue simulation and interactive storytelling. These transformations have significant implications for both authors and readers, particularly with regard to the reshaping the notions of authorship and creativity in relation to authenticity. The presence of AI in literary creation redefines who or what can be considered an author and challenges readers’ expectations about the origin and intentionality of a text (Grace and Sahu, 2024).

In light of these recent developments, relatively few studies in the field of literature and literary education have explored the processes of generation and reception of AI-generated texts. Of those that do exist, most have focused on the general public or adult readers. Relatively few studies (e.g., Saddhono et al., 2024) focus on adolescent audiences, despite the fact that young people are increasingly engaged with digital ecosystems.

This paper addresses this gap by shifting the focus to adolescents, who are a particularly relevant group for whom the distinction between human-authored and AI-generated poetry is influenced by their recent literary education in school and a high level of familiarity with digital technologies. Having been exposed to canonical poetry through the national curriculum while simultaneously interacting with environments that normalise algorithmic creativity, adolescents are not “typical” readers, but a diagnostically relevant group for examining how the distinction between human-authored and AI-generated poetry is perceived and processed. Their interpretive practices, shaped both by structured literary frameworks and by familiarity with technology-mediated authorship, differ from those of adult readers and provide insights into how future generations of readers may perceive and react to the distinction between human and machine creativity.

The study is based on a few research questions concerning mainly the emotional and creativity reception of AI-generated poetry among adolescents. More specifically, it asks whether AI-generated poetry can evoke authentic emotional responses as human-written poetry and how “label effects” of authorship influence poetic evaluation. It also examines how adolescents respond when the origin of the text is initially unknown and how these responses change once it is disclosed that the poem was generated by AI, thereby testing beliefs about authenticity. To address these questions, I proposed a case study based on two key variables: (1) a sample of 12th-grade adolescents who were recently engaged in Romanian literary education, and (2) the assessment of emotional impact and perceived creativity and complexity across the two types of texts. The final aim is to compare the reception of AI-generated and human-written poems and to analyse the effect of author disclosure. Investigating how factors such as gender and educational profile influence the reception of poetry remains a secondary objective.

As a point of departure, it is assumed that adolescents may respond more emotionally to AI-generated poetry as its language may resonate more with their everyday speech. However, evaluations are expected to continue being shaped by label and “authority effects” related to author identity as synthesised in the literature reviewed by Nabasirye (2024: 14–18), particularly when texts are associated with canonical authors. Adolescents’ reception may therefore remain influenced by patterned interpretative frameworks which are developed through school-based literary practices, especially in Romania. I will refer to this phenomenon as ‘patterned reception’ hereafter.¹

2 Toward a new poetry: from AI-generated texts to literary reception

The relationship between literary creation, reading practices and technological mediation has been debated for a few decades, long before the appearance of LLMs. Earlier discussions focused on how digital environments reshaped reading processes, machine-assisting reading, the beginnings of generative texts through algorithmic processes. For instance, Hayles (2012) introduced the concept of “technogenesis,” referring to the co-evolution of humans and technology, to analyse how digital environments reshape human cognition and reading practices, including the shift from deep reading to “hyperreading” and “machine reading.” Understanding this co-evolution of humans and technology is essential because it reveals the continuity between earlier technological influences on literary practices and reading habits and current interactions with AI-generated texts. Rather than constituting a radical rupture, LLMs extend long-standing patterns through which literary development, reading practices and cognitive engagements with language have continually adapted to technological change. The release of advanced

¹ Although the syntagm “patterned reception” may appear in different research areas, in the present study it is used to reflect a local, context-specific reality, referring to the structured ways in which Romanian adolescents evaluate literary texts through pre-established interpretative frameworks rather than text-internal features. These patterns are formed through repeated exposure to school-based literary commentary and locally situated educational practices.

AI models such as ChatGPT (GPT-4) in 2022 has led to a significant transformation in paradigms of literary generation and reception, giving rise to new processes of textual production and meaning-making, as readers are now confronted not only with technologically mediated texts, but with texts whose production is completely automated, leading to a reconfiguration of authorship and aesthetic value.

This AI-driven redefinition of literature is particularly important today because it helps us recognise that we may be moving toward entirely new forms of expression shaped by AI, which does more than automatically generate text. Nowadays, AI can also attempt to approximate the subtle “human touch” of literary creation, especially the metaphorical depth that derives from human cognition. However, there are some recent studies (Amirjalili et al., 2024) that highlight LLMs’ limitations in reproducing both the distinctive cognitive imprint of a human author, particularly the context-sensitive use of metaphor, and the rigor of academic style – at least at this moment. As the authors argue, AI still shows shortcomings in depth, expressivity, and preserving the distinctive attitude of the author’s voice and personal human touch.

Against this theoretical background, it is necessary to ask how these dynamics are currently unfolding in the field of literature as AI has begun to assume an increasingly active role in the creation and co-creation of texts. AI can rapidly generate narrative or poetic patterns, which authors can later modify. Even before ChatGPT’s release, there were authors who claimed that their texts were guided or even primarily produced by AI² – some of which have been adapted for theatre and popularized around the world.³

When it comes to poetry, a range of dedicated “poetry generation” models has emerged over time. Linardaki (2022) traces the evolution of this field from early experiments to the creation of increasingly complex poetic compositions. From a theoretical standpoint, AI-generated poetry seems to be primarily a combination of words and stylistic patterns – an approach that has also been explored during previous transitions between literary movements, but without technological mediation. In the present, however, this process takes the form of a technologically mediated textual production that combines existing poetic patterns on a large scale as AI poetry generators⁴ are trained on extensive literary corpora that can produce poems that closely resemble those written by human authors. Nevertheless, the significance of these poetry generators lies more in their ability to reshape reader expectations and the value that readers assign to poetic texts than in their ability to produce texts.

Besides these automatic text generators, there are also perspectives that emphasize the positive influence of AI on literary creativity and they should not be neglected. Selvi and Ramya (2024), in their study

Application of AI in Literature: A Study on Evolution of Stories and Novels, argue that AI “simplifies and accelerates the writing process by supporting idea generation, plot development, and character creation. It also helps overcome writer’s block, enables the exploration of new narrative directions and tailors stories to audience preferences. These technologies democratize storytelling, making it more accessible and interactive” (Selvi and Ramya, 2024). Overall, these perspectives emphasise that AI can serve as a co-creative tool, enhancing the writing process. This is particularly evident in its ability to help authors overcome creative obstacles, as highlighted by Selvi and Ramya, while also increasing the efficiency of the writing process by reducing the time required for idea generation and narrative development. At the same time, while these models offer valuable support in generating ideas and modifying form, ongoing concerns about authenticity and textual integrity remain central to the debate.

Beyond questions of literary generation and co-creation, a key topic is how readers receive, interpret and evaluate AI-generated poetic texts. As the role of the human author in the creation of texts becomes less significant, earlier theories of literary reception that anticipated the “death of the author” can be revisited. For instance, in his theory of reception, Barthes (1967) interprets this “death” metaphorically, arguing that meaning is not embedded by the author but constructed by the reader. In the context of AI-generated texts, however, this idea takes on a more literal dimension, as human authorial intervention can be greatly reduced. Consequently, when the author is no longer present in the traditional sense, or when his role is redefined, attention could shift even more strongly toward the impact on the reader and the processes of reception of the literary texts.

In this regard, the contemporary cognitive theories of reading retains elements of this reader-centred perspective, emphasising that interpretation relies on the dynamic construction of “mental models” based on textual hints and the reader’s prior knowledge (Kintsch, 1998). Given the metaphorical nature of human mind, readers navigate between concrete and abstract conceptual domains (Lakoff and Johnson, 1980), while mental representations amplify the emotional experience of the text (Stockwell, 2002). Thus, literary reception results from a complex interaction between text, context and reader, shaped by affective and cognitive engagement (Zwaan, 2004). This framework of literary reception is even more relevant when authorship is technologically mediated, as it invites reflection on how AI-generated texts reshape reading habits and expectations, influence perceptions and even inform decisions about what human creativity means.

In the last few years, a growing number of studies have examined AI-generated poetry from the perspective of literary reception, focusing on how readers perceive and evaluate such texts. A notable contribution to this field is the study by Hitsuwari et al. (2023), which investigates how readers evaluate poems created or mediated by AI, depending on the degree of human intervention involved in the creation process. In this study, the authors compared three categories of haiku: those written by humans, those generated by AI without human input (“human-out-of-the-loop”), and those generated with human intervention (“human-in-the-loop”). The study found that “human-in-the-loop” poems received the most favourable evaluations, while those fully generated by AI were rated lower, potentially reflecting a bias against AI authorship – an issue also explored by Bellaiche et al. (2023). Another relevant study focused on reader experience that influenced this research belongs to Porter and Machery (2024), which examined the reception of AI-generated

2 See “Death of an Author,” written under the guidance of Canadian Stephen Marche, who published it under the pseudonym Aidan Marchine.

3 For example, the play “A.I.: When a Robot Writes a Play,” directed by Daniel Hrbek, director of the Švanda Theatre in Prague, is considered the first theatre play written by AI. Premiered in 2021, it has been performed on many international stages, including in Bucharest, where I must admit I saw it myself at the beginning of 2025. More information can be found here: <https://www.svandovodivadlo.cz/en/inscenations/673/ai-when-a-robot-writes-a-play>.

4 For example, see AI-generated poetry tools such as <https://www.aipoemgenerator.org/> and <https://deepai.org/chat/poet>.

poetry among over 1,400 American participants. Their findings also revealed a preference for AI-generated poetry, likely due to their simpler and more accessible language. In contrast, the complexity of human-written poetry was often perceived as vague or inaccessible, a phenomenon they defined “more human than human” (Porter and Machery, 2024).

As LLMs continue to improve in semantic and pragmatic dimensions – including the recognition of secondary meanings, figurative language, irony, and humour – the role of AI in literary creation is likely to strengthen. Moreover, this development raises a pertinent question about the reception of AI-generated literary texts. If a text produced via simple prompts evokes the same emotions, semantic depth, and intentionality as one crafted by a human being, to what extent is the human author still relevant in this creative process? Are we at a crossroads in defining human creativity, given that minimally edited AI-generated texts may outperform those written by humans in reader evaluations? It is crucial to assess the impact of such texts on readers and whether they might diminish readers’ critical capacity to interpret intent, emotion, and linguistic nuances.

3 The methodological approach and sample characteristics

For the present study, poems by four of the Romanian authors included in the national school curriculum were selected: Mihai Eminescu, Lucian Blaga, Nichita Stănescu, and Mircea Cărtărescu.

These authors were selected due to their familiarity among adolescents. Each of them is also representative of a different literary movement. Three of the authors – Mihai Eminescu, Lucian Blaga, and Nichita Stănescu – are included in the official syllabus for the national baccalaureate examination. The fourth author, Mircea Cărtărescu, is a prominent contemporary postmodernist writer, who was included to also examine potential differences in reception related to postmodernism as a literary movement and author contemporaneity. Furthermore, including his poetry in the questionnaire facilitated an exploration of how contemporary relevance influences the reception of poetry among adolescents. This is particularly relevant given that his work is nowadays the subject of ongoing literary debate.

As previously mentioned, the selected Romanian writers represent four distinct literary movements within Romanian literature, namely romanticism (Mihai Eminescu), modernism/expressionism (Lucian Blaga), neo-modernism (Nichita Stănescu), and postmodernism (Mircea Cărtărescu). This selection ensures a high level of stylistic and thematic diversity, thereby supporting the study’s comparative objectives related to poetry reception.

One representative poem per author was selected for inclusion in the questionnaire given to participants. The poems were drawn from literary volumes that are referenced in Romanian language and literature textbooks for grades 11 and 12, ensuring a shared curricular context and comparability between authentic literary texts and AI-generated texts in terms of thematic focus and stylistic particularities.

In addition, a literary corpora of 20 texts per author was selected from the same volumes and used as in-context prompts to guide the chosen AI model for this study – ChatGPT-4o, in

generating new texts. This approach ensured coherence between the original poems included in the questionnaire and the AI-generated texts.

3.1 AI-stimulus generation procedure

For each author, ChatGPT was provided with structured background information, including biographical context, key thematic and conceptual features, and relevant critical references. This data was sourced from two officially approved textbooks: the Romanian Language and Literature Textbook, grade XI, authored by Dobra et al. (2006), and grade XII, coordinated by Simion et al. (2007), both endorsed by the Ministry of Education and Research, by OME no. 4446 of 19.06.2006, respectively OMEdCT no. 1783/2 of 16.08.2007. The selection of textbook-based content in the prompt design was intentional, as it facilitated the assessment of officially endorsed educational knowledge.

To generate the AI poem, an iterative prompting procedure was employed using ChatGPT-4o for each of the four Romanian authors. Interaction with ChatGPT-4o took place in colloquial language typical of conversation, conveying that a study on the reception of AI-created poetry was underway and that support was required for its completion. In this context, the model was provided with preliminary information indicating that it would receive data on four Romanian authors from different literary movements, structured as follows for each of them: (1) biographical information, (2) fundamental concepts of their work, (3) critical references and (4) a set of 20 representative literary texts for each of the four authors for processing. The task was to *generate a new poetic text that resemble each author’s style in terms of thematic concerns, structure, and symbolism*. The process involved five successive prompt iterations per author, each progressively refined through targeted instructions concerning thematic depth, distinctive literary style and symbolism, detailing to key motifs, tone, language register, and prosody, where applicable. For example, when generating a text in the style of Mihai Eminescu, ChatGPT-4o was instructed progressively to refine the language register to reflect 19th-century usage and Romantic prosody, whereas for the version attributed to Mircea Cărtărescu it was prompted to incorporate postmodern, colloquial, and fragmented modes of expression. The fifth iteration was selected due to its qualitative improvements following the cumulative feedback given throughout the previous iterations and because it provided a consistent and methodologically appropriate stopping point for comparative analysis across all texts. The full prompt wording, author-specific refinements, and details of each iteration are provided in [Supplementary Appendix 8](#).

3.2 Poem matching criteria (human vs. AI)

After ChatGPT-4o generated the four poetic texts, each AI-generated poem was subsequently paired with a corresponding human-authored text by the same author, selected to ensure maximum comparability. The matching process was carried out manually and guided by three main criteria: (a) thematic and symbolic coherence – including shared motifs or conceptual focus and (b) structural and formal similarity – aiming for comparable length and prosodic structure.

3.3 Participant recruitment and assignment

To examine the reception of poetic texts among adolescents, including both AI-generated and human-authored poems, a target group of 100 adolescents aged 17–19 was selected. Participants were recruited from four 12th grade high school classes within the same theoretical high school. Participation was voluntary, and all students were invited to take part in the study after being informed and providing consent. The classes were randomly assigned to the control and experimental groups, with each group comprising an equal number of students ($n = 50$ per group). Since the participants belonged to intact classroom units, potential class-level clustering was considered and addressed in the interpretation of results, though no statistical adjustment was applied due to the small number of clusters.

The selected adolescents came from four distinct academic profiles (1) mathematics and computer science, (2) philology, (3) social sciences, and (4) natural sciences. The gender distribution in the control group was 66% female and 34% male, whereas in the experimental group it was 64% female and 36% male. Although the gender distribution was not perfectly balanced between groups, the presence of both genders in each group allowed for an overall analysis of perceptual differences between females and males. As the study involved human participants, it was conducted in accordance with established ethical research standards, all responses being anonymized to ensure confidentiality. The study posed no psychological or physical risks to those involved.

The two groups (control and experimental) were managed differently in terms of questionnaire administration. Participants in the control group were informed that they would analyse eight poems, four of which had been generated by an AI model. Participants were asked to read each poem and respond to questions about their appreciation of the poems, their emotional impact and their opinion of the creativity and complexity of the texts. Participants in the experimental group were presented with the same set of poems and questions, but without any information about the authorship. The aim here was to evaluate their immediate, uninfluenced reactions to the literary texts. After answering the same set of questions as the control group, the experimental group was informed that the final four poems had been generated by an AI model. They were then given the opportunity to re-read these texts and re-evaluate them.

Data were collected using a structured questionnaire divided into two sections. The first section contained questions aimed at identifying the demographic characteristics of the respondents, including their age and academic profile. The second section comprised four, respectively five questions related to the reception of literary texts, as follows:

1. Participants were asked to rate how much they like each poem on a scale from 1 (not at all) to 5 (very much).
2. The second question aimed to assess the intensity of the emotional response evoked by each poem. Participants were asked: “*What emotion did the poem evoke in you, on a scale from 1 (not at all) to 5 (very strongly)?*” A table was provided with eight emotional categories: the six basic emotions identified by Darwin (1872) and Ekman and Friesen (1975) – joy, sadness, anger, fear, surprise, and disgust – as well as “no emotion” and “other emotion.” Participants could rate each

emotion individually on a five-point Likert scale, allowing for multiple emotional reactions to be recorded for the same text.

3. The third question assessed the extent to which the poem was perceived as creative, using a scale from 1 to 5.
4. The fourth question involved evaluating the language complexity of the poem, again on a five-point scale.
5. The fifth question (for the experimental group only) followed the disclosure that the final four poems had been generated using artificial intelligence (ChatGPT-4o). Participants were asked whether they wished to revise any of their previous responses. This was a binary question (‘yes’ or ‘no’); if they answered ‘yes,’ participants were invited to provide revised answers for the four AI-generated texts. This approach was chosen to facilitate a comparative analysis between spontaneous reactions and those influenced by awareness of the poems’ AI-generated origin. The aim was to assess the impact of this knowledge on participants’ evaluations.

4 Analysing and interpreting results

Analyses of participants’ responses were conducted using descriptive statistical methods, with a focus on the comparative presentation of means and variances of responses, using SPSS. Given that the study was designed as a pilot investigation aimed at identifying patterns of reception, inferential statistical testing was not employed as the primary analytical strategy. Accordingly, observed variations between groups, text types, or evaluation stages are reported only as tendencies within the sample, rather than as statistically generalized effects. To address potential concerns regarding subgroup differences and uneven group sizes, a limited inferential strategy was adopted at the beginning of the analysis, alongside the study’s primarily descriptive orientation. Repeated measures of General Linear Model (GLM) analyses were employed, nevertheless, as the sole inferential approach. The purpose of the tests was to assess whether the descriptive pattern observed in the data was consistent with interaction effects between the author and the experimental condition (control vs. experimental), as well as the two main respondent characteristics, gender and academic track. The above-mentioned analyses are in place as a preliminary robustness check, rather than as the primary basis for interpretation. The intention was to evaluate if aggregate descriptive trends masked substantial interaction structures under conditions of unequal subgroup sizes and repeated measurements. After this initial inferential step, all subsequent analyses and interpretations are presented descriptively. Inferential terminology is deliberately avoided in order to maintain the emphasis on descriptive and pattern identification. In addition, linguistic features were examined in an illustrative manner, with the aim of contextualising readers’ emotional responses rather than conducting a systematic text analysis.

To assess adolescents’ reactions to the texts, I analysed the responses for each poem presented to both groups, calculating the means and standard deviations for all assessed questions. This approach aimed to capture both the major trends and the degree of variation (dispersion) within each group.⁵

⁵ The full data sets are presented in [Supplementary Appendices 1–6](#).

TABLE 1 Correlation of texts for comparative analysis.

Text No.	Author name	Correlated with Text No.	AI model prompted on the author's literary references and corpora	Group analysed	Analysis conducted
Text 1	Mihai Eminescu	Text 5	Mihai Eminescu	Control/ Experimental	Calculate mean, standard deviation, score difference
Text 2	Lucian Blaga	Text 6	Lucian Blaga	Control/ Experimental	Calculate mean, standard deviation, score difference
Text 3	Nichita Stănescu	Text 7	Nichita Stănescu	Control/ Experimental	Calculate mean, standard deviation, score difference
Text 4	Mircea Cărtărescu	Text 8	Mircea Cărtărescu	Control/ Experimental	Calculate mean, standard deviation, score difference

As mentioned before, each original text written by a human author was matched with the AI-generated counterpart text obtained by prompting the AI model on that author's work, biography and critical references. Thus, the texts⁶ were analysed in pairs corresponding to the four targeted authors (1–5, 2–6, 3–7, 4–8, as shown in Table 1). For each pair, I calculated the mean score for each question in the questionnaire, separately for the control and experimental groups, as well as the mean difference in scores between the human and AI-generated poetry. This reflects how the poems was appreciated and received according to the degree of respondents' information about the author of the texts.

The results of the preliminary repeated-measures GLM analyses conducted are fully listed in Appendices 9 and 10. In the light of the pattern-orientated design of the study, these models were employed in order to fulfil robustness checks rather than to serve as the main basis for interpretation. Data analysed interactions between author, experimental condition and respondent characteristic. Both tables contain multivariate tests (Pillai's Trace) within-subjects and polynomial contrast for each evaluative and emotional dimension. Gender moderation delivered limited and outcome-specific results that are notable, especially in the case of general appreciation and disgust.

Gender moderation outlines that the effects vary across emotions and do not strongly differentiate responses to AI-generated versus human-authored poetry based on gender. In contrast to gender, educational profile shows clearer evidence of selective moderation. This is the case for general appreciation, sadness, fear, surprise, disgust or creativity. Data suggest that author-specific variation relies rather on distinct evaluative orientation when responding to certain textual features based on educational profile. While the moderation effects underscore that they are not pervasive, these are rather confined to some specific aspects.

These preliminary interaction checks indicate to us that while subgroup characteristics such as gender or educational profile could shape responses in selected and bounded ways, these account less for the overall pattern of reception in the data. In this sense, it is highlighted that the analysis is based on descriptive statistics between AI-generated and human-authored poems, emphasising general tendencies in evaluation and emotional response rather than subgroup-specific effects.

Based on the correlation between the texts written by Romanian authors and those generated by AI, a comparative analysis of the adolescents' answers was carried out. The following Table 2 shows the mean differences in evaluation between poems written by Romanian

authors and those generated by AI, according to whether the respondents belonged to the control or the experimental group. Positive values in all the following tables indicate a preference for poems written by human authors, while negative values reflect a positive evaluation of AI-generated texts. Correlating these values with variables such as the respondents' belonging to a particular group (control or experimental) revealed relevant trends in adolescents' perceptions of poetry. This correlation also formed the basis for establishing perceptual links between the particularities of the authors and their literary movements, and the emotional responses and degrees of appreciation expressed by the respondents.

The results show that, when readers are not aware of the author of a text, differences in perception and emotional response between the two categories of texts are reduced and, in some cases, reversed. In such situations, AI-generated texts receive higher evaluations than those written by human authors. As discussed earlier, this phenomenon can be described in terms of 'patterned reception', at least in local context. Thus, respondents in the experimental group, who were not informed about the author of the texts at the outset, were more likely to rate AI-generated poems more highly and often evaluated them more favourably than poems written by human authors. By contrast, in the control group, where the adolescents knew the author of the texts from the outset, higher evaluations were recorded for texts attributed to Romanian authors. This pattern may be discussed in relation to the recognised literary status of the selected authors, who are representative figures of Romanian literature.

For instance, the difference in overall preference for original and AI-generated poetry for Lucian Blaga changes from +0.16 in the control group to -0.54 in the experimental group, reflecting a clear reversal of preference in favour of the artificially generated text when the author is unknown (see Table 2). This reversal can be read as illustrating the effect of the absence of authorial information on the poem's reception. The effect is also observable in responses to questions about creativity and complexity, with similar shifts recorded, pointing to a more open and possibly more honest evaluation of AI-generated poems. A similar trend is recorded for Mircea Cărtărescu, where the difference in the overall rating changes from +0.48 in the control group to -0.24 in the experimental group, again reflecting higher evaluations for AI-generated poetry. The differences in the evaluations of human and AI-generated poetry are smaller in terms of the overall rating for Mihai Eminescu and Nichita Stănescu; for instance, the difference in the overall rating drops only from +0.78 (control group) to +0.02 (experimental group) for Mihai Eminescu.

In terms of emotions, poems written by human authors are more frequently associated with positive emotions, especially when the author is known. This pattern can be read as indicating that authorial labelling is associated not only with aesthetic evaluations but also with

⁶ The eight texts used in the questionnaires are included in Supplementary Appendix 7.

TABLE 2 Mean difference between the ratings of original and AI-generated poems by author and respondent group.

	Control group				Experimental group			
	Mihai Eminescu	Lucian Blaga	Nichita Stănescu	Mircea Cărtărescu	Mihai Eminescu	Lucian Blaga	Nichita Stănescu	Mircea Cărtărescu
General appreciation	0,78	0,16	0,32	0,48	0,02	-0,54	0,06	-0,24
Happiness	0,79	0,71	0,56	0,35	0,02	0,32	0,66	0,32
Sadness	0,68	-0,61	-0,92	-0,02	-0,10	-0,73	-0,87	-0,07
Anger	-0,39	-0,26	-0,10	-0,08	-0,34	-0,37	-0,34	-0,23
Fear	-0,07	-1,01	-0,39	-0,10	-0,30	-0,61	-0,31	-0,27
Surprise	0,33	-0,09	0,19	0,82	-0,03	0,02	0,51	0,42
Disgust	-0,36	-0,25	-0,26	-0,08	-0,07	-0,24	-0,12	0,17
No emotion	-0,91	0,27	-0,11	0,06	0,02	0,25	0,10	0,17
Creativity	0,80	0,38	0,58	0,80	-0,20	-0,42	0,06	0,20
Complexity	0,94	0,84	0,22	0,96	-0,02	0,08	-0,26	0,34

Values represent mean difference scores (Original – AI). The control group comprised $N = 50$ participants (female = 33, male = 17; sciences = 25, humanities = 25). Valid sample sizes vary across emotion measures within the control group: general appreciation ($N = 50$), happiness ($N = 31-43$), sadness ($N = 37-43$), anger ($N = 27-30$), fear ($N = 28-30$), surprise ($N = 29-32$), disgust ($N = 21-28$), no emotion ($N = 24-26$), creativity ($N = 50$), and complexity ($N = 50$). The experimental group comprised $N = 50$ participants (female = 32, male = 18; sciences = 16, humanities = 34). Valid sample sizes within the experimental group were as follows: general appreciation ($N = 50$), happiness ($N = 30-35$), sadness ($N = 31-41$), anger ($N = 26-32$), fear ($N = 26-32$), surprise ($N = 32-36$), disgust ($N = 26-30$), no emotion ($N = 26-28$), creativity ($N = 50$), and complexity ($N = 50$).

emotional responses to poetry. In contrast, AI-generated poems are more frequently associated with negative emotions, such as sadness or disgust. For instance, for Lucian Blaga and Nichita Stănescu, negative sadness scores are consistent in both the control (-0.61 ; -0.92) and experimental (-0.73 ; -0.87) groups. Furthermore, fear is rated more intensely in AI-generated poems, particularly for Lucian Blaga (-1.01 in the control group), with higher fear-related scores recorded for AI-generated text. By contrast, positive emotions such as happiness are more frequently associated with human poetry. In this respect, Mihai Eminescu's poems recorded a value of $+0.79$ in the control group, while this value is much lower in the experimental group.

Knowing the author is associated with differences in perceptions of the creativity and complexity of poems. Adolescents record higher creativity and complexity-related evaluations for human poems when they know the author, a pattern that may be discussed in relation to the impact of the author's identity and prestige on literary reception. This is evident in the cases of Mihai Eminescu ($+0.80$ for creativity and $+0.94$ for complexity in the control group), Lucian Blaga ($+0.38$ for creativity and $+0.84$ for complexity) and Mircea Cărtărescu ($+0.80$ for creativity and $+0.96$ for complexity). However, in the absence of knowledge about the author, these differences are highly reduced or no longer observed. For example, Mihai Eminescu's texts receive lower creativity and complexity-related scores when the author was unknown (-0.20 and -0.02). This pattern shows that AI texts receive comparable creativity and expressivity-related evaluations when authorship information is absent. A special case is Nichita Stănescu, for whom creativity is evaluated similarly for both types of text, although AI poetry is rated as less complex (-0.26).

Based on these general data, I next calculated the mean, standard deviation and interquartile range (IQR) for each group of respondents, presented in Table 3, in order to describe the distribution of responses, in particular by examining dispersion around the median values (through the IQR). At this stage, I focused on describing the consistency of evaluations within each group, to assess the extent to which responses were divergent or homogeneous and to compare

these variations between the control and experimental groups, regardless of the author, focusing only on the respondent groups.

Overall, the negative mean in the experimental group (-0.18) reflects higher evaluations for AI poems over human poems when the author is unknown. This pattern may be read as corresponding to a form of reception that is less constrained by authorial labelling and less shaped by prior expectations, with evaluations appearing more neutral overall. In comparison, the positive mean recorded for the control group (0.44) reflects higher evaluations for human poetry over AI-generated poetry when the author is known. In both cases, the standard deviation was moderate and similar, indicating that responses were relatively homogeneous and not characterised by wide dispersion.

Regarding emotions, respondents in both the control and experimental groups associated sadness particularly with AI-generated poems. The assessment showed high IQR values for both groups (0.843 and 0.673), indicating substantial variability in responses. This variability is observable in relation to the reception of AI texts associated with Nichita Stănescu (control group: -0.92 ; experimental group: -0.87) and Lucian Blaga (control group: -0.61 ; experimental group: -0.73). These values are noticeably higher than those of the other two authors. Regardless of whether respondents knew the origin of the texts, they also identified anger and fear as emotions experienced after reading. AI-generated poems are associated with higher anger-related scores (-0.21 in the control group and -0.32 in the experimental group). There was high agreement (IQR = 0.198) when the respondents knew the author of the poem and near-unanimity (IQR = 0.035) when they did not.

Several interpretative perspectives regarding the higher negative emotion-related scores recorded for AI-generated poems, such as sadness or anger, may be discussed. One such perspective relates to the current limitations of the LLMs, which have been described as producing texts with reduced nuances compared to human language. In this context, AI-generated poems may be evaluated as more rigid or emotionally harsher. Similarly, the lower degree of metaphorical

TABLE 3 Mean, standard deviation and interquartile range in the evaluation of original and AI-generated poems, regardless of author, by respondent group.

	Control group			Experimental group		
	Average	Standard deviation	IQR	Average	Standard deviation	IQR
General appreciation	0,44	0,265	0,275	-0,18	0,277	0,345
Happiness	0,60	0,193	0,223	0,33	0,262	0,160
Sadness	-0,22	0,705	0,843	-0,44	0,417	0,673
Anger	-0,21	0,146	0,198	-0,32	0,062	0,035
Fear	-0,39	0,436	0,453	-0,37	0,159	0,093
Surprise	0,31	0,381	0,333	0,23	0,275	0,435
Disgust	-0,24	0,116	0,078	-0,06	0,172	0,140
No emotion	-0,17	0,516	0,423	0,14	0,098	0,110
Creativity	0,64	0,202	0,270	-0,09	0,275	0,350
Complexity	0,74	0,351	0,260	0,03	0,248	0,225

Values represent mean difference scores (Original – AI), collapsed across all four authors. Positive values indicate higher ratings for original poems, whereas negative values indicate higher ratings for AI-generated poems. The control group comprised $N = 50$ participants (female = 33, male = 17; sciences = 25, humanities = 25), and the experimental group comprised $N = 50$ participants (female = 32, male = 18; sciences = 16, humanities = 34). Valid sample sizes vary across emotion measures due to item-level missing responses; all means are computed using available data for each variable.

ambiguity characteristic of human-authored poetry may be reflected in higher negative emotion-related evaluations. Additionally, prompting the AI model on texts that explore deeper themes concerning human existence may result in the creation of literary structures and images that intensify feelings of anxiety and outrage. Finally, it cannot be excluded that, particularly in the control group, some respondents may deliberately associate AI-generated poems with negative emotional registers.

At the same time, AI poems are associated with higher reported fear ratings in both the control group (-0.39) and the experimental group (-0.37). Analysing the IQR values reveals that, when the origin of the text is unknown, responses cluster more closely, with adolescents recording higher fear-related ratings (IQR = 0.093). This pattern can be described in relation to recurring language features observed in the AI texts, which include a higher frequency of negatively marked lexical units and constructions. For example, only as an illustrative text-based comparison, Lucian Blaga's original text contains no verb accompanied by negation, whereas, seven verbs in negative constructions appear in the text generated by the model prompted on this author's poetry: 'You will not find me', 'I will not die', 'do not see', 'do not ask', 'I would not have left', 'it does not set', 'you will not see her again'. Additionally, within the same text, lexical units and structures associated with semantic fields linked to loss or death are present, such as 'I will be buried', 'stone or clay', 'the thrill of the wind', 'the forgotten dawn', 'shadow', 'it is setting', 'deeper silence', and so on. Similar lexical patterns are also noticed in other AI-generated texts, where recurrent lexical choices associated with negative semantic fields appear. Across these texts, such patterns are observable at multiple linguistic levels, particularly lexical and semantic, through the repeated presence of negatively marked semantic constructions, the use of adjectives occurring in dysphoric semantic contexts, and other lexical units or structures referring to absence, burden, and death.

Adolescents reported higher sadness ratings when reading texts generated by AI than when reading texts written by human authors. However, this difference appears less pronounced than that observed for other emotions. As mentioned above, sadness ratings following AI

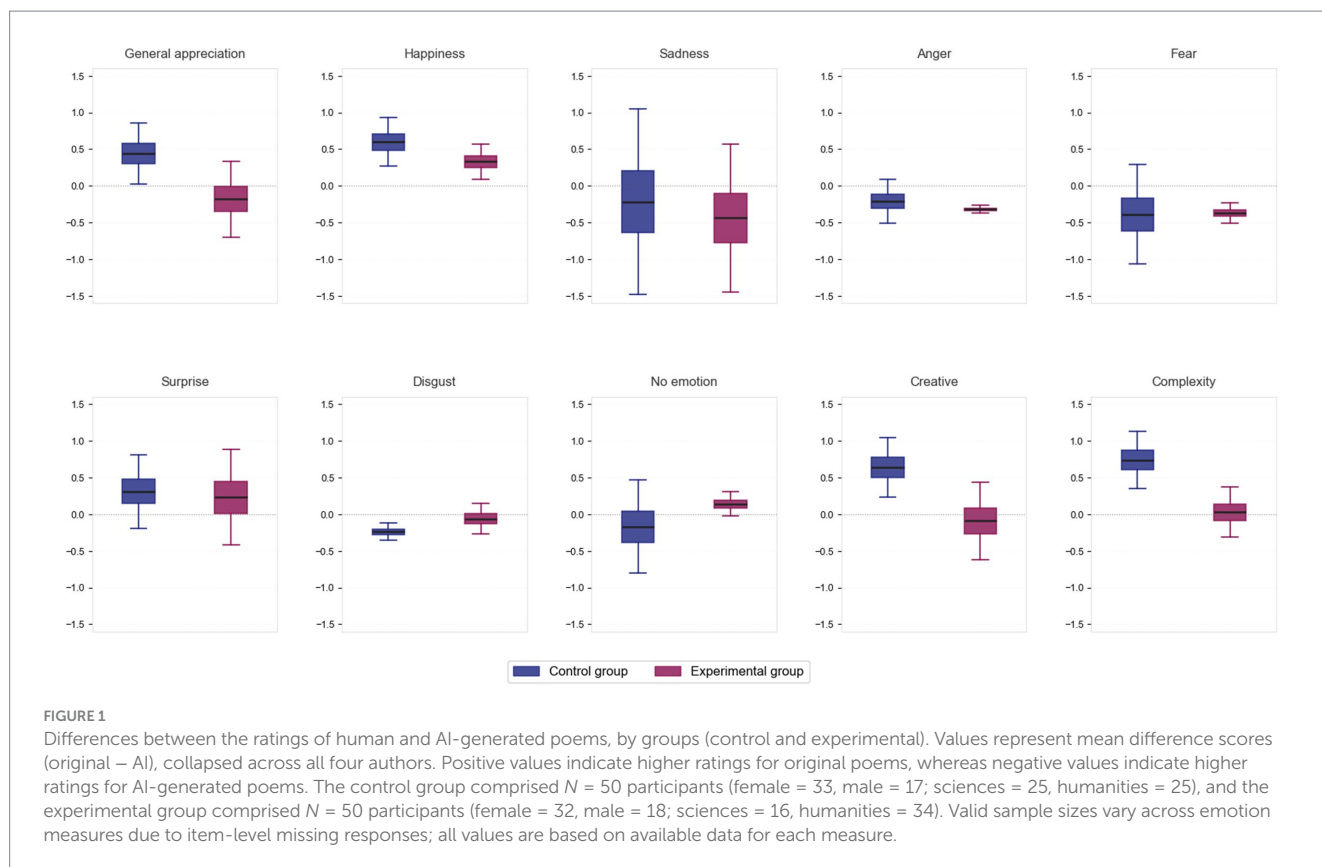
texts were higher in both the control group (-0.22) and the experimental group (-0.44). However, the results appear less consistent in comparison with those of other analysed emotions, as the high IQR values recorded in the control (0.843) and experimental (0.673) groups indicate major variations among respondents, i.e., a lack of clear consensus on the intensity of sadness reported.

This variation is observable alongside differences in respondents evaluations across authors, co-occurring with differences in familiarity with or expectations of each author. For instance, the mean difference in evaluation between original and AI-generated poems was -0.92 in the case of Nichita Stănescu and 0.68 in the case of Mihai Eminescu. In these cases, differences in evaluation are noticed in relation to author-related expectations.

The graph below (Figure 1) illustrates the effect of authorship on evaluations of human and AI-generated poems for each analysed dimension (general appreciation, emotional response, creativity, and complexity). Generally, there is a substantial change in perception between the control and experimental groups, especially regarding sadness, creativity and complexity.

Knowledge of the origins of a poem appears to intensify negative perceptions of AI-generated texts. In the experimental group, for example, the mean score difference was small (-0.06 , IQR = 0.140), which corresponds to a relatively neutral distribution of responses. By contrast, the mean score registered a decrease in the control group (-0.24 , IQR = 0.078), reflecting lower average ratings for AI poem in this group. This difference between the two groups may be discussed in relation to current prejudices about AI, given that it represents an artificial language model and is often associated with shallowness and 'depersonalisation'.

Among participants who stated that the texts did not evoke any emotions, ratings indicating a lack of emotional response were more frequently assigned to poems written by human authors than to AI poems, particularly when the respondents were unaware of the text's origin. In the experimental group, the mean score for this dimension was positive (0.14), whereas in the control group it was negative (-0.17). This pattern corresponds to lower average ratings of



emotional evocativeness for AI-generated poems when respondents were aware of the texts' origin. The relatively low IQR value (0.110) in the experimental group shows homogeneity of response. The largest difference in recorded values in the perception of human versus AI-generated poetry was observed in Mihai Eminescu's text. In the control group, the mean difference in evaluation was -0.91 ; in the experimental group, however, it was close to zero (0.02). These values describe higher average emotional ratings for AI-generated poems in the absence of information about authorship. An exception is Lucian Blaga's text, for which the aggregated evaluation values remained close to zero even when information about the author was available. Given the fact that across several dimensions responses displayed internal consistency within classroom groups, this can also be read as indicative of a clustering effect at the classroom level.

The perceived level of creativity also varies in relation to respondents' knowledge of the author's identity. In the control group, for example, the overall mean was positive ($+0.64$), corresponding to a higher average creativity ratings for poems written by human authors. By comparison, the mean in the experimental group was negative but close to zero (-0.09), reflecting more similar average ratings for human and AI-generated texts. However, the moderate IQR values recorded in both groups -0.270 in the control group and 0.350 in the experimental group – indicate relative heterogeneity in the responses. For instance, in the experimental group, AI-generated texts based on the work of Mihai Eminescu and Lucian Blaga were rated as more creative than the originals, suggesting that less complex and hermeticised AI poetic language may be perceived as more persuasive to adolescents. But in a broader sense, a descriptive analysis of the respondents' answers in this case emphasises that recognizing

the origin of a poem is linked to how its creativity is evaluated. When readers were aware that a text was generated by AI than rather than written by a human, it receives lower evaluations of creativity.

Similarly, knowledge of the author of the text is associated with differences in the evaluation of the degree of complexity of poetic language. In both the control group (mean = 0.74) and the experimental group (mean = 0.03), texts written by human authors were generally rated as more complex than those generated by artificial intelligence. At the same time, except for Mircea Cărtărescu's original text, which received higher complexity ratings in both groups, respondents did not differentiate between human-generated and AI-generated texts. Furthermore, in some instances, AI-generated poems were considered more complex, a pattern observable in the recorded evaluations across texts.

To summarise the results obtained from evaluating the original and AI-generated poems, Table 4 shows the score differences between the two types of text based on a conventional system of signs (“+,” “++,” “+++,” “0,” “-,” “--,” “---”). These symbols reflect respondents' relative evaluations of one type of author over another: a “+” indicates a more positive evaluation of human poetry, while a “-” indicates a higher appreciation of AI-generated texts. The “0” symbol signals a minimal difference between the two. According to the coding convention used:

- “+++” marks a strongly positive difference in favour of human text (above $+0.75$)
- “++” a moderately positive difference ($+0.35$ to $+0.75$)
- “+” a slightly positive difference ($+0.10$ to $+0.34$)
- “0” a minimal difference (-0.09 to $+0.09$)
- “-” a slightly negative difference (-0.10 to -0.34)

- “--” a moderately negative difference (−0.35 to −0.75)
- “---” a strongly negative difference (below −0.75) in favour of AI texts.

Also, to summarise the impact of authorial identity on literary reception exclusively, Figure 2 shows that, in the control group, human texts are generally considered more expressive and complex and receive higher overall evaluations. By contrast, AI-generated poems are more frequently rated as neutral, or in some cases are associated with intense negative emotions such as sadness or fear.

At the end of the questionnaire, respondents in the experimental group were asked to re-evaluate the AI-generated poems using the same set of questions they had previously answered (see Table 5). In the post-evaluation phase, there were marked decreases in the scores for overall appreciation (−0.64), creativity (−0.69), complexity (−0.42) and sadness (−0.47) compared to the mean scores for these characteristics of each of the four authors. These changes are observed following the disclosure of non-human authorship and correspond to lower post-evaluation scores. This pattern may be discussed in relation to existing biases associated with AI, as well as scepticism and fear regarding its creative potential.

Adolescents were even more surprised (+0.39) and happier (+0.14) when they found out that the texts were generated by AI. This may be interpreted in relation to adolescents’ excitement about the technology’s ability to produce such convincing poetry. Regarding negative emotions such as anger (−0.06), fear (−0.20) and disgust (+0.11), similar perceptions were evident post-assessment, with minimal differences, suggesting consistency in the recorded evaluations of these negative affective dimensions.

However, the high standard deviation values at the post-assessment stage indicate a lack of consensus among participants regarding the AI-generated poems, with meaningful variation in their responses. Some adolescents re-evaluated the AI texts more critically, while others adopted a neutral stance. This dispersion of responses may be discussed in relation to individual differences, including personal beliefs, educational profile, gender and perhaps

even the level of familiarity with technology of the respondents, as we will see later.

The decrease in scores at the post-assessment stage corresponds to higher recorded evaluations for poems written by human authors compared to those generated by AI. Knowledge of the text’s artificial origin is associated with lower evaluations of expressiveness, particularly in relation to creativity and complexity.

The following graph (Figure 2) summarises and compares the differences in evaluation between poems written by human authors and those generated by AI, according to the author and whether the respondents belong to the control or experimental group. Accordingly, the direction and intensity of the evaluation (in favour of human or AI poetry) are represented through an intuitive chromatic scale: green for favourable evaluations of human texts and red for favourable evaluations of AI-generated texts. Several recurring patterns can be observed, such as the experimental group’s consistently higher recorded evaluations of Lucian Blaga’s AI-generated poetry, and the decrease in positive evaluations of Mihai Eminescu’s original poetry when the author is unknown.

5 The influence of personal characteristics on the reception of poetry

5.1 Respondents’ educational profile

To analyse the extent to which the educational profile of adolescents is associated with differences in their perception of poems, the mean answer for each question was calculated according to profile, group of respondents, and author. The two major academic tracks – the humanities (philology and social sciences) and the sciences (mathematics and computer science, and natural sciences) – were used as independent variables to assess differences in poetry reception, in terms of the emotions evoked as well as general appreciation, and perceptions of the creativity and complexity of the texts.

TABLE 4 Comparative summary of evaluations of human vs. AI poetry by authors and groups.

	Control group				Experimental group			
	Mihai Eminescu	Lucian Blaga	Nichita Stnescu	Mircea Cărtărescu	Mihai Eminescu	Lucian Blaga	Nichita Stnescu	Mircea Cărtărescu
General appreciation	+++	+	++	++	0	–	0	–
Happiness	+++	+++	++	++	0	+	++	+
Sadness	++	–	–	0	–	–	–	–
Fury	–	–	–	–	–	–	–	–
Fear	–	–	–	–	–	–	–	–
Surprise	+	–	+	+++	0	0	++	++
Disgust	–	–	–	–	–	–	–	+
No emotion	–	+	–	0	0	+	0	+
Creativity	+++	+	++	+++	–	–	0	+
Complexity	+++	+++	+	+++	0	0	–	+

Symbols indicate the direction and magnitude of mean difference scores (Original – AI). Cutoff values were derived from the distribution of difference scores across all measures and authors. Sample composition was as follows: control group, N = 50 (female = 33, male = 17; sciences = 25, humanities = 25) and experimental group, N = 50 (female = 32, male = 18; sciences = 16, humanities = 34). Valid sample sizes vary across emotion measures due to item-level missing responses.

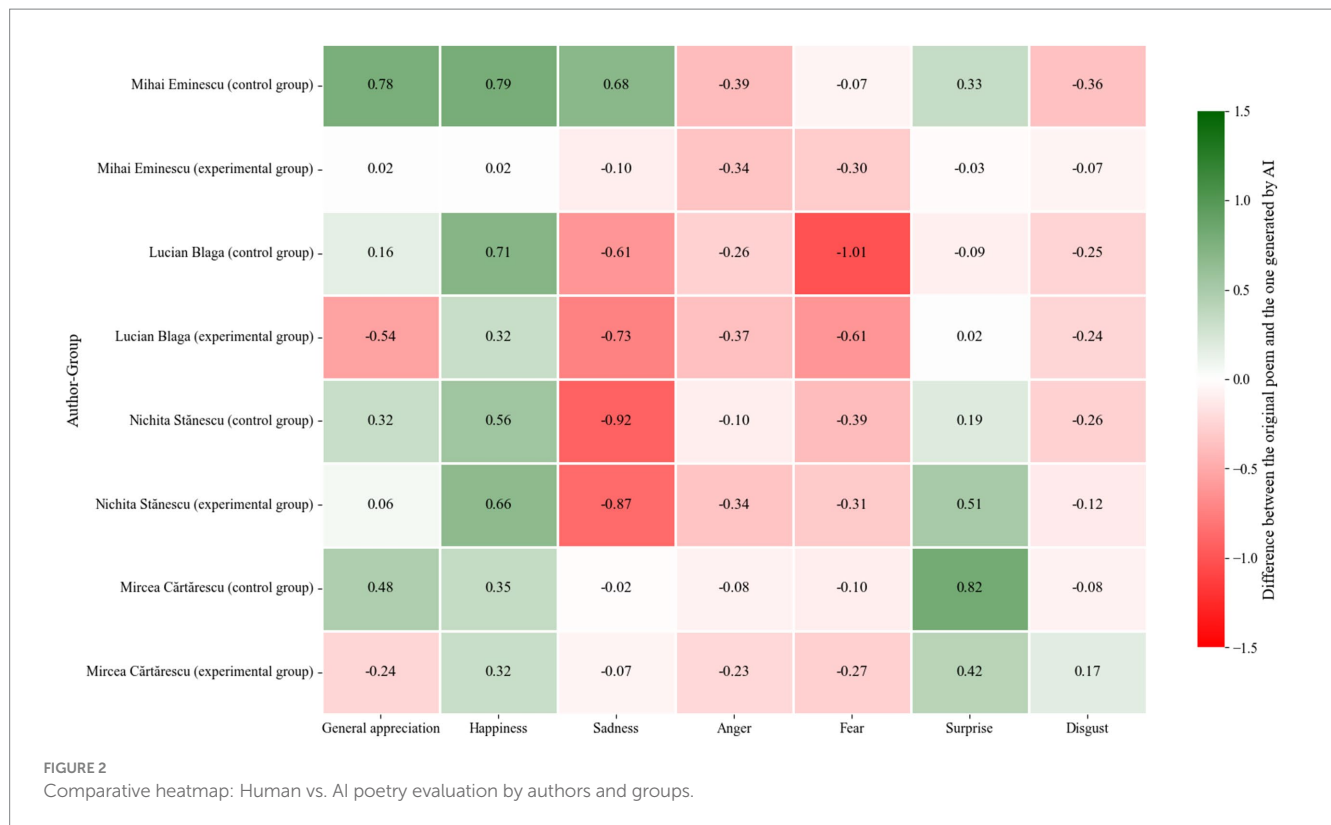


TABLE 5 Mean ratings recorded for AI poems in the experimental group (pre/post text source knowledge).

	Average		Standard deviation	
	Experimental group	Post-evaluation	Experimental group	Post-evaluation
General appreciation	3,82	3,18	0,873	1,190
Happiness	2,16	2,30	0,999	1,102
Sadness	2,89	2,42	0,834	1,001
Fury	1,49	1,43	0,700	0,774
Fear	1,85	1,65	0,990	0,950
Surprise	2,29	2,68	1,005	1,358
Disgust	1,22	1,33	0,529	0,802
No emotion	1,74	1,66	1,164	1,010
Creativity	3,87	3,18	0,716	1,119
Complexity	3,66	3,24	0,806	1,170

Values represent mean ratings (1–5 scale) of AI-generated poems, shown separately for the initial blind evaluation and the post-disclosure re-evaluation. The experimental group comprised N = 50 participants (female = 32, male = 18; sciences = 16, humanities = 34). Valid sample sizes vary due to item-level nonresponse: general appreciation, creativity, and complexity (N = 50); happiness (N = 30–35); sadness (N = 31–41); anger and fear (N = 26–32); surprise (N = 32–36); disgust (N = 26–30); and no emotion (N = 26–28). Post-disclosure re-evaluation sample sizes were as follows: general appreciation, creativity, and complexity (N = 50); happiness (N = 37); sadness (N = 33); anger (N = 30); fear (N = 34); surprise (N = 38); disgust (N = 30); no emotion (N = 29); and other emotion (N = 33).

The data calculated below (Table 6) were subsequently aggregated according to the different measured aspects. The raw data used to calculate the above two tables can be found in Supplementary Appendix 4. This approach was used to capture patterns in how an upper secondary education relates to literary reception. For instance, adolescents specialising in the humanities are exposed to greater amounts of reading and literature due to the provisions in the educational framework plans. These provisions include more hours of studying for language and literature, as well as

a wide range of optional subjects focusing on national and world literature and reading. This context is relevant when interpreting differences observed between educational profiles, without inferring differential analytical ability compared to adolescents specialising in sciences. In what follows, I will highlight the differences observed between the two educational profiles, emphasising where they converge or differ in their assessment of poetry.

The results from the experimental group show that, when information about the author of the texts is absent, adolescents

TABLE 6 Mean differences in emotions between original and AI-generated poems by respondent group (control/experimental) and educational profile.

	Sciences		Humanities		Δ Sciences - Humanities	
	Control group	Experimental group	Control group	Experimental group	Control group	Experimental group
General appreciation	0,14	-0,54	0,26	-0,56	-0,13	0,02
Happiness	0,44	0,57	0,84	0,42	-0,41	0,15
Sadness	-0,49	-0,46	-0,47	-0,95	-0,02	0,49
Anger	-0,12	-0,16	-0,39	-0,29	0,28	0,13
Fear	-0,12	0,10	-0,21	-0,31	0,09	0,41
Surprise	0,06	0,37	0,14	-0,25	-0,07	0,63
Disgust	0,34	0,04	-0,27	0,11	0,61	-0,07
No emotion	-0,26	0,45	-0,22	0,21	-0,04	0,24
Creativity	0,64	-0,05	0,64	-0,11	0,00	0,06
Complexity	0,48	0,00	0,67	-0,05	-0,19	0,05

Values represent mean difference scores (Original – AI), aggregated across all four authors. Δ Sciences – Humanities was calculated as (science-track mean – humanities-track mean). Subgroup sizes were as follows: control group—sciences ($N = 25$), humanities ($N = 25$); experimental group—sciences ($N = 16$), humanities ($N = 34$). Valid sample sizes across emotion measures (both groups combined) were: general appreciation, creativity, and complexity ($N = 50$); happiness ($N = 37$); sadness ($N = 33$); anger ($N = 30$); fear ($N = 34$); surprise ($N = 38$); disgust ($N = 30$); no emotion ($N = 29$); and other emotion ($N = 33$).

assigned higher average scores to AI-generated poems than to poems written by human authors, regardless of their educational profile (sciences: -0.54 , humanities: -0.56). For instance, in the experimental group, adolescents with a background in the humanities gave higher scores to AI-generated poems based on the work of Mihai Eminescu (-0.03) and Mircea Cărtărescu (-0.50) than to the original texts. Similarly, adolescents with a background in sciences gave higher scores to AI-generated poems based on the work of Lucian Blaga (-2.19) and Nichita Stănescu (-0.43). In this case, the difference between the two profiles is almost non-existent, at just 0.02 points.

In the control group, adolescents with a background in the humanities gave poems written by Romanian authors an average score that was 0.26 points higher than scores given to poems generated by AI. However, those with a background in science only gave poems written by Romanian authors an average score that was 0.14 points higher than AI-generated poems. These values indicate higher average evaluations of poems written by Romanian authors when information about authorship was available.

This pattern is more evident among those with a humanities background, who assign higher scores to Romanian authors' texts. In the experimental group, both categories of adolescent assigned higher average scores to AI texts. In the control group, however, higher evaluations of human texts are observed in the presence of expectations associated with the prestige status of Romanian authors, regardless of the respondents' educational background. These expectations may be discussed in relation to the school environment in which the adolescents are immersed during their time as students.

In terms of emotional reception, poems written by human authors received higher happiness-related scores, regardless of the respondents' educational profile. Differences between original and AI-generated texts are more pronounced when adolescents know the origin of the poem. This can be understood in relation to the role of author status in the reception of literary texts, as well as to a patterned reception. Interestingly, in the experimental group, adolescents with a background in science assigned higher average scores to the original

texts of three of the four authors than their peers with a background in the humanities. An exception is Nichita Stănescu's text, for which this trend is not maintained – a notable aspect. It is also worth mentioning that positive emotions received higher evaluations when postmodern, contemporary authors are referred to, as with Mircea Cărtărescu's text. This observation may be interpreted as indicating a preference for a poetic language closer to colloquial speech among adolescents.

As previously mentioned, sadness is rated as a more intense emotion after reading AI-generated texts, particularly when the poem's origin is unknown to the respondents. This tendency is more pronounced among humanities adolescents (mean: -0.95) than among science adolescents (mean: -0.46). A notable example of this is the reception of the AI text attributed to Lucian Blaga, which received the highest sadness-related scores in both the experimental group (sciences: -1.67 ; humanities: -2.36) and the control group (sciences: -1.73 ; humanities: -1.54). Earlier in this section, linguistic markers possibly evoking this emotion were presented as illustrative examples. Nevertheless, it is still relevant to observe differences between the two categories of adolescents in the experimental group, where those with a background in humanities report higher sadness ratings to AI-generated poems than their peers from sciences.

In general, AI's ability to express sadness appears more plausible for poets whose work is characterised by introspection or melancholic language. For instance, the differences between the original and AI-generated texts of Nichita Stănescu and Mihai Eminescu were relatively moderate. However, in the case of Mircea Cărtărescu, these differences were minimal, with closely aligned evaluation scores, reflecting a limited differentiation between human and AI-generated texts. This similarity in perception may be viewed in relation to the particularities of postmodern poetic discourse, which is characterised by colloquial, self-referential, and less evocative language, features that are frequently observable in AI-generated texts.

Regarding anger, it can be observed that it is not a dominant theme in any of the analysed poems, and is only associated with

slightly higher scores in the case of the AI-generated texts. However, some specific tendencies are noticeable and worth emphasising. Firstly, knowledge of the poem's origin is associated with differences in evaluation, depending on their educational background. In the case of the poem attributed to Mihai Eminescu, respondents with an academic background in sciences rated lower anger-related scores when they knew the text was generated by AI, whereas those studying humanities assigned higher scores for this dimension. By contrast, this situation is reversed in the case of Nichita Stănescu's poetry. There are also considerable variations for the other authors according to the high school profile, indicating that evaluations of anger vary across authors and respondent profiles, without implying a single explanatory factor.

Fear, one of the most prominent emotions evoked in general, does not vary substantially between categories of respondents who know the poem's origin, regardless of their profile. However, those studying humanities report higher fear-related ratings when reading an AI-generated poem than those studying sciences, regardless of whether they are in the control or experimental group. In the case of Lucian Blaga's poem, adolescents enrolled in sciences profile who know the origin of the poem do not distinguish between human and AI-generated poetry. However, in the experimental group, these same respondents recorded higher fear-related scores when evaluating the human text. On the other hand, adolescents studying humanities in the control group reported higher levels of fear in Nichita Stănescu's original poem than in the AI-generated text. In the case of Mircea Cărtărescu's original poem, there are no major differences in the perception of fear among respondents with a background in sciences. However, knowledge of the text's origin is associated with lower fear-related scores in the control group. Therefore, when respondents are informed that a poem has been generated by AI, lower fear-related evaluations are recorded and the distinction between human-written and AI-generated text is reduced, except for adolescents studying the humanities.

Regarding the evaluation of disgust according to educational profile, there are no major differences, irrespective of whether respondents belong to the control or experimental group. However, some variations can be noted, which may be discussed in relation to stereotyping or respondents' expectations. In the case of adolescents studying humanities, for example, knowledge that the poetry is AI-generated is associated with higher disgust-related ratings, particularly in the case of poetry attributed to Mihai Eminescu. This pattern may be interpreted in relation to a rejection of the idea that AI can reproduce the work of an author such as "Mihai Eminescu, the national poet of Romania." A spectacular increase in disgust is observed among respondents studying sciences when the poem is attributed to Lucian Blaga, with the mean difference in recorded values increasing from 0.23 in the experimental group to 1.60 in the control group. This heightened level of disgust may be discussed in relation to expectations associated with the philosophical dimension of Lucian Blaga's poetry, as opposed to an artificial text perceived as lacking that sacred, philosophical substance. This reaction can also be interpreted as a form of symbolic 'profanation' of the author's values.

Respondents studying humanities record higher disgust-related evaluations for AI-generated poems when they realise the artificial origin of the texts. This tendency is evident in the case of most authors. For AI poems generated according to the poems of Mihai Eminescu, for example, the difference in perception increases from -0.12 in the

experimental group to -0.61 in the control group. For Nichita Stănescu, it increases from -0.18 to -0.63 . In the case of Mircea Cărtărescu, there is a reversal of attitude: an initially favourable evaluation of the AI text of $+0.30$ becomes a negative evaluation of -0.09 once the AI authorship is identified.

AI-generated poetry is associated with a range of reactions among adolescents, regardless of their educational background, especially when they know the texts are AI-generated. Thus, AI texts based on Mihai Eminescu's romantic poetry or Lucian Blaga's expressionist poetry received lower surprise-related scores than those inspired by Nichita Stănescu or Mircea Cărtărescu. The observed trends are relatively similar for the two educational profiles. However, adolescents specialising in science subjects recorded higher surprise-related scores for the poem attributed to Mircea Cărtărescu, whereas those specialising in humanities showed higher surprise-related evaluations for the text associated with Nichita Stănescu. This pattern may be discussed in relation to the literary periods of the analysed poets. While the poetic language of Mihai Eminescu and Lucian Blaga adheres more closely to traditional stylistic norms, such as well-defined prosodic elements like rhyme and metre, the transition to neomodernism (Nichita Stănescu) and postmodernism (Mircea Cărtărescu) is characterised by increasingly self-referential, colloquial and flexible poetic language, with fewer poetic norms. These linguistic features may facilitate a closer alignment between AI-generated and human-written texts, which is associated with higher levels of reported surprise among adolescents.

Furthermore, adolescents who initially stated that the texts did not evoke any emotion show changes in their responses when aware of the artificial origin of the texts, with higher evaluations recorded for AI-generated poems. Among respondents with a background in sciences, the differences are considerable. Regarding Mihai Eminescu's poetry, for example, the evaluation ranges from moderate appreciation of the original text (0.46 in the experimental group) to strong rejection of the AI-generated text (-1.00 in the control group). Similar variations are seen with Lucian Blaga (from 1.57 to 0.13), though smaller differences are seen with Nichita Stănescu (0.10 to 0.00) and Mircea Cărtărescu (-0.35 to -0.16).

The proportion of adolescents who say that AI-generated poems do not evoke any emotion is higher when they are aware that the poems were created using AI. The greatest variations are seen in poems attributed to Mihai Eminescu and Lucian Blaga, followed by Nichita Stănescu and Mircea Cărtărescu. Notably, respondents with a humanistic profile record lower emotional evaluations for the original poetry by Mircea Cărtărescu, even when they know the author. This observation may be seen in relation to the postmodern poetic language, which is more fragmented and colloquial and thus closer to the language used by young people, a stylistic context that may be associated with reduced emotional engagement.

Regarding the perception of text creativity according to educational profile, respondents unanimously assign higher creativity ratings to works written by human authors than to AI texts when they know the origin of the poetry. Lucian Blaga is the only author whose original work was received lower creativity-related scores than that generated by AI by both sciences and humanities profile in the experimental group. Also, among the humanities respondents who were unaware of the poem's source, Mihai Eminescu's text was rated as less creative than its AI counterpart (-0.29), whereas the sciences respondents did not record notable differences in the creativity of the

two types of text. A similar situation was found in the evaluations of adolescents with an academic track in sciences, who assigned lower creativity ratings to Nichita Stănescu's poem than the AI-generated counterpart (-0.12). Without contextual information about authorship, higher evaluations of innovation and expressiveness are recorded for AI-generated texts, regardless of educational profile (Table 7).

The perception of the complexity of human poetry varies in relation to knowledge of the text's origin, with a more pronounced variation observed among humanities respondents. For instance, the recorded complexity ratings of Mihai Eminescu's poetry increase from -0.06 to 0.12 among adolescents with a background in sciences, whereas among those with a humanities profile, the variation is much wider, ranging from -0.41 to 0.44 . Similar patterns are also evident in the evaluation of Lucian Blaga's work (sciences profile: from -0.19 to 0.80 ; humanities profile: from 0.21 to 0.88). Also, Nichita Stănescu's original poetry receives lower complexity ratings than AI-generated poetry among sciences adolescents, regardless of their knowledge of the source. By contrast, Mircea Cărtărescu is the only author whose original text is rated as more complex than the AI-generated counterpart, regardless of educational profile or prior knowledge of the poems' source. Notably, humanities respondents rate Mircea Cărtărescu's poetry as less creative than those with a background in sciences. It is important to note that the differences associated with educational profile also coincide with the classroom-based organisation of the sample, indicating that some patterns may reflect shared evaluative frameworks within classes.

5.2 Gender of the respondents

To highlight gender differences in terms of general appreciation, emotional reception and perception of the creativity and complexity of the analysed texts, I calculated the mean response for each item according to gender, group and author. The resulting values are presented in Table 8 and formed the basis for further thematic aggregation by dimension of analysis, as shown in Table 9. The full dataset used to construct these tables is available in Supplementary Appendix 5.

Male respondents recorded higher scores for negative emotions generated by AI poems, such as sadness, anger, fear, and disgust. This pattern may be discussed in relation to differences in emotional response to AI-generated texts, even in the absence of a human author. This is particularly prominent among adolescents in the experimental group, where the origin of the poems was unknown, a context in which higher emotion-related scores were recorded.

Conversely, female respondents consistently rated poems written by human authors more favourably than male respondents, but especially in the control group. At the same time, female respondents showed a more critical attitude in their evaluations of AI texts when their origin was known, reflected in lower overall appreciation scores. Additionally, these differences in perception between the two genders may be interpreted as reflecting different modes of sensitivity to poetic texts, as expressed through the evaluation scores.

The emotional response according to the gender of the respondents emphasises that happiness-related evaluations are more frequently associated with poetry written by human authors, regardless of their gender. Noticeable differences appear in the case of

Mihai Eminescu's and Lucian Blaga's texts, where male respondents record higher evaluations of their poems in both the control and experimental groups. In this case, female respondents record lower evaluations of poems written by human authors and assign relatively higher evaluations to AI poetry, especially when they do not know the author.

Regarding sadness, the differences in perception according to gender are more pronounced in the case of Lucian Blaga and Nichita Stănescu. In the experimental group, both male (Blaga: -1.22 , Stănescu: -1.08) and female (Blaga: -0.58 , Stănescu: -0.98) respondents reported higher sadness-related scores for AI-generated poems, but there were important differences between the two genders. Moderate differences between genres were observed in the case of Mihai Eminescu, where the general tendency was to associate sadness with his original poems. For Mircea Cărtărescu, female respondents recorded higher sadness-related scores for the original poem when they do not know the author – and for the AI-generated text in the control group. Therefore, the more pronounced differences in sadness-related evaluations among male respondents may be discussed in relation to different evaluative responses to poems of unknown origin, without inferring emotional authenticity judgments.

The evaluation of anger is higher in the case of AI-generated poems, but knowledge of the poem's origin is associated with lower anger-related scores among male respondents, regardless of the author – Mihai Eminescu ($-0.52 \rightarrow -0.27$), Lucian Blaga ($-0.55 \rightarrow 0.08$), Nichita Stănescu ($-0.64 \rightarrow -0.09$), Mircea Cărtărescu ($-0.34 \rightarrow -0.09$). In the case of female respondents, this pattern is less pronounced. Thus, in the control group, anger is more pronounced in the case of poems written by Mihai Eminescu and Lucian Blaga, remaining relatively constant in the case of Nichita Stănescu and Mircea Cărtărescu.

On a general level, fear is more frequently associated in the recorded evaluations with AI-generated poems. Similar to anger, reception also varies according to gender. For male respondents, knowledge of the author is associated with lower fear-related scores for AI texts. In contrast, for female respondents, fear is rated more highly when they know the text is AI-generated. The exception is Lucian Blaga, where the score in the experimental group (-0.61) is less negative than in the control group (-1.01).

Interestingly, surprise is more frequently associated with human poetry overall. However, scores vary depending on the gender of the reader and their knowledge of the author. Among female respondents, there is a decrease in surprise for the poetry of Lucian Blaga (from 0.25 to -0.19) and Nichita Stănescu (from 0.58 to 0.18), when the origin of the text is known. In the case of Nichita Stănescu, on the other hand, the scores associated with AI poetry increase in both gender groups, while those associated with human poetry decrease. The highest levels of surprise are recorded for Mircea Cărtărescu's poetry, followed by Mihai Eminescu's.

Regarding disgust, it can be observed that it is weakly evoked, but differentiated according to gender. Once again, knowledge of the author is associated with lower disgust-related scores in the case of males ($-0.32 \rightarrow -0.17$), but with higher disgust-related scores in the case of females ($0.08 \rightarrow -0.29$), to the detriment of AI.

Broader considerations of emotional reception show that AI-generated poems are more frequently associated with a lack of emotion, a phenomenon that is more pronounced among male respondents. The greatest differences are found in the male control

TABLE 7 Mean differences in emotions, by author, experimental or control group and educational profile, between original and AI-generated poems.

		Sciences		Humanities	
		Control group	Experimental group	Control group	Experimental group
Mihai Eminescu	General appreciation	0,48	0,13	1,08	-0,03
	Happiness	0,56	0,29	1,10	-0,10
	Sadness	0,76	0,00	0,59	-0,15
	Anger	-0,28	-0,40	-0,51	-0,30
	Fear	0,05	-0,24	-0,24	-0,33
	Surprise	0,17	0,50	0,54	-0,32
	Disgust	-0,19	0,00	-0,61	-0,12
	No emotion	-1,00	0,46	-0,80	-0,25
	Creativity	0,84	0,00	0,76	-0,29
	Complexity	0,12	-0,06	0,44	-0,41
Lucian Blaga	General appreciation	-1,02	-2,19	-0,56	-2,00
	Happiness	0,43	0,92	1,17	0,84
	Sadness	-1,73	-1,67	-1,54	-2,36
	Anger	-0,20	-0,06	-0,62	-0,09
	Fear	0,06	0,64	-0,25	0,00
	Surprise	-0,98	-0,92	-1,00	-1,22
	Disgust	1,60	0,23	0,25	0,43
	No emotion	0,13	1,57	-0,19	0,50
	Creativity	0,40	-0,50	0,36	-0,38
	Complexity	0,80	-0,19	0,88	0,21
Nichita Stănescu	General appreciation	0,44	-0,43	0,20	0,29
	Happiness	0,31	0,49	0,85	0,73
	Sadness	-0,85	-0,53	-1,03	-1,03
	Anger	-0,14	0,09	-0,05	-0,55
	Fear	-0,42	0,18	-0,34	-0,57
	Surprise	0,04	1,16	0,39	0,28
	Disgust	-0,01	0,00	-0,63	-0,18
	No emotion	0,00	0,10	-0,21	0,11
	Creativity	0,56	-0,12	0,60	0,15
	Complexity	-0,12	-0,38	0,56	-0,21
Mircea Cărtărescu	General appreciation	0,64	0,31	0,32	-0,50
	Happiness	0,44	0,56	0,25	0,21
	Sadness	-0,13	0,36	0,10	-0,26
	Anger	0,15	-0,27	-0,39	-0,22
	Fear	-0,16	-0,18	0,00	-0,33
	Surprise	1,01	0,75	0,61	0,25
	Disgust	-0,06	-0,09	-0,09	0,30
	No emotion	-0,16	-0,35	0,32	0,47
	Creativity	0,76	0,44	0,84	0,08
	Complexity	1,12	0,63	0,80	0,20

Values represent mean difference scores (Original – AI) reported separately for each author. Subgroup sizes were as follows: control group—sciences (N = 25), humanities (N = 25); experimental group—sciences (N = 16), humanities (N = 34). Valid sample sizes across emotion measures (total sample) were: general appreciation, creativity, and complexity (N = 50); happiness (N = 37); sadness (N = 33); anger (N = 30); fear (N = 34); surprise (N = 38); disgust (N = 30); no emotion (N = 29); and other emotion (N = 33). These totals are distributed across the 16 author × condition × educational-track combinations displayed in the table.

TABLE 8 Mean differences in emotions, by author, experimental or control group and gender, between original and AI-generated poems.

		Male		Female	
		Control group	Experimental group	Control group	Experimental group
Mihai Eminescu	General appreciation	0,48	0,13	1,08	-0,03
	Happiness	0,56	0,29	1,10	-0,10
	Sadness	0,76	0,00	0,59	-0,15
	Anger	-0,28	-0,40	-0,51	-0,30
	Fear	0,05	-0,24	-0,24	-0,33
	Surprise	0,17	0,50	0,54	-0,32
	Disgust	-0,19	0,00	-0,61	-0,12
	No emotion	-1,00	0,46	-0,80	-0,25
	Creativity	0,84	0,00	0,76	-0,29
	Complexity	0,12	-0,06	0,44	-0,41
Lucian Blaga	General appreciation	-1,02	-2,19	-0,56	-2,00
	Happiness	0,43	0,92	1,17	0,84
	Sadness	-1,73	-1,67	-1,54	-2,36
	Anger	-0,20	-0,06	-0,62	-0,09
	Fear	0,06	0,64	-0,25	0,00
	Surprise	-0,98	-0,92	-1,00	-1,22
	Disgust	1,60	0,23	0,25	0,43
	No emotion	0,13	1,57	-0,19	0,50
	Creativity	0,40	-0,50	0,36	-0,38
	Complexity	0,80	-0,19	0,88	0,21
Nichita Stănescu	General appreciation	0,44	-0,43	0,20	0,29
	Happiness	0,31	0,49	0,85	0,73
	Sadness	-0,85	-0,53	-1,03	-1,03
	Anger	-0,14	0,09	-0,05	-0,55
	Fear	-0,42	0,18	-0,34	-0,57
	Surprise	0,04	1,16	0,39	0,28
	Disgust	-0,01	0,00	-0,63	-0,18
	No emotion	0,00	0,10	-0,21	0,11
	Creativity	0,56	-0,12	0,60	0,15
	Complexity	-0,12	-0,38	0,56	-0,21
Mircea Cărtărescu	General appreciation	0,64	0,31	0,32	-0,50
	Happiness	0,44	0,56	0,25	0,21
	Sadness	-0,13	0,36	0,10	-0,26
	Anger	0,15	-0,27	-0,39	-0,22
	Fear	-0,16	-0,18	0,00	-0,33
	Surprise	1,01	0,75	0,61	0,25
	Disgust	-0,06	-0,09	-0,09	0,30
	No emotion	-0,16	-0,35	0,32	0,47
	Creativity	0,76	0,44	0,84	0,08
	Complexity	1,12	0,63	0,80	0,20

Values represent mean difference scores (Original – AI) reported separately for each author. Positive values indicate higher ratings for original poems, whereas negative values indicate higher ratings for AI-generated poems. Subgroup sizes were as follows: control group—female ($N = 33$), male ($N = 17$); experimental group—female ($N = 32$), male ($N = 18$). Total group sizes were $N = 50$ for both the control and experimental groups. Valid sample sizes across emotion measures (total sample) were: general appreciation, creativity, and complexity ($N = 50$); happiness ($N = 37$); sadness ($N = 33$); anger ($N = 30$); fear ($N = 34$); surprise ($N = 38$); disgust ($N = 30$); no emotion ($N = 29$); and other emotion ($N = 33$). These totals are distributed across the 16 author \times condition \times gender combinations shown in the table.

TABLE 9 Mean differences in emotions overall, by experimental or control group and gender of respondents, between original and AI-generated poems.

	Female		Male		Δ Female - Male	
	Control group	Experimental group	Control group	Experimental group	Control group	Experimental group
General appreciation	0,36	-0,27	0,59	-0,02	-0,23	-0,25
Happiness	0,58	0,25	0,68	0,49	-0,10	-0,24
Sadness	-0,18	-0,29	-0,32	-0,70	0,14	0,41
Anger	-0,28	-0,22	-0,09	-0,51	-0,19	0,30
Fear	-0,39	-0,24	-0,42	-0,59	0,03	0,35
Surprise	0,31	0,34	0,37	0,05	-0,07	0,29
Disgust	-0,29	0,08	-0,17	-0,32	-0,12	0,39
No emotion	-0,07	0,33	-0,35	-0,16	0,28	0,48
Creativity	0,62	-0,16	0,69	0,03	-0,07	-0,19
Complexity	0,75	0,03	0,72	0,06	0,03	-0,03

Values represent mean difference scores (Original - AI), aggregated across all four authors. Δ Female - Male was calculated as (female mean - male mean). Subgroup sizes were as follows: control group—female ($N = 33$), male ($N = 17$); experimental group—female ($N = 32$), male ($N = 18$). Total group sizes were $N = 50$ for both the control and experimental groups. Valid sample sizes across emotion measures (total sample) were: general appreciation, creativity, and complexity ($N = 50$); happiness ($N = 37$); sadness ($N = 33$); anger ($N = 30$); fear ($N = 34$); surprise ($N = 38$); disgust ($N = 30$); no emotion ($N = 29$); and other emotion ($N = 33$).

group (Eminescu: -1.18 ; Blaga: -0.09 ; Cărtărescu: -0.31). Even in the experimental group, male respondents record higher scores indicating an absence of emotion in AI-generated poetry (Eminescu: -0.67 ; Blaga: -0.09 ; Stănescu: -0.05), whereas female respondents gave higher ratings to AI-generated poetry (Eminescu and Blaga: $+0.45$; Stănescu: $+0.24$). Notably, female respondents record lower emotional evaluations for Mircea Cărtărescu's original poetry than for the AI-generated text, in contrast to the male respondents' pattern of associating AI-generated poetry with lower emotional content.

Creativity remains a dimension of reception in which human poetry receives higher evaluations from all respondents, regardless of gender. For females, scores increase for human poetry between the two groups: Mihai Eminescu ($+1.01$), Mircea Cărtărescu ($+0.82$), Nichita Stănescu ($+0.77$), Lucian Blaga ($+0.49$). In the case of male respondents, the largest differences also occur for Lucian Blaga ($+1.37$) and Mihai Eminescu ($+1.00$), while for Mircea Cărtărescu ($+0.19$) and Nichita Stănescu ($+0.08$), the variations are moderate. However, in the experimental group, both female and male respondents rated the AI-generated poetry of Eminescu and Blaga as more creative. Mircea Cărtărescu is the only author whose original poetry received higher creativity-related scores than AI-generated poetry, regardless of gender.

The perception of a poem as complex shifts in favour of human poetry when respondents know its origin, regardless of gender. But in the experimental group, male (-0.17) and female (-0.31) respondents both assigned higher complexity ratings to Nichita Stănescu's AI poetry. Male respondents also assigned higher complexity ratings to AI poetry in the case of Mihai Eminescu. However, in the control group, where the human author was known, evaluations consistently favoured human poems, with balanced differences between genders – as in the case of Mihai Eminescu's poems.

In summary, the gender of the respondent is associated with differences in the evaluation of AI-generated poetry, particularly when they do not know the author of the texts (see Table 9). Female respondents display a more differentiated pattern of emotional

evaluations, with lower scores for certain negative emotions (e.g., anger) compared to male respondents. By contrast, the most substantial differences emerged in responses concerning the absence of emotional reactions ('no emotion'). Female respondents in the experimental group recorded lower scores for emotional absence in AI poems than male respondents ($+0.48\%$). It is worth noting that some of the gender-related differences observed also overlap with the classroom structure of the sample, suggesting that part of this variation can be attributed at the level of shared classroom contexts.

6 Conclusion

It may appear surprising that adolescents liked and rated AI-generated poetry more favourably than poetry written by human authors, especially when they did not know the origin of the texts. In some cases, AI poems were considered to be more creative or received higher emotion-related evaluations than poems written by human authors. However, once the origin of the texts was revealed, the experimental group recorded lower evaluations of AI poems, with changes observed across all analysed dimensions: overall appreciation, emotional reaction, creativity and complexity. The largest decreases were recorded for the AI texts associated with Mihai Eminescu and Lucian Blaga. This pattern of re-evaluation can be interpreted in the context of prevailing stereotypes about AI-generated poetry and points to broader questions about how readers interpret non-human poetic expression.

The emotional response to a poem varies in relation to whether it is written by a human or an AI model. When the author is known, AI-generated poems receive lower emotion-related evaluations than human poems. However, when this information is absent, human poems sometimes receive lower emotion-related evaluations than AI-generated poems. Negative emotions such as sadness and disgust are more often associated with AI-generated texts, particularly when the author's poetic language is characterised by philosophical or

introspective features (e.g., Lucian Blaga and Nichita Stănescu). On the other hand, happiness is more often associated with human poems, especially when the author of the texts is known.

When respondents know the author, creativity and complexity are more frequently associated in the recorded evaluations with human poetry. In the control group, knowing the author is associated with higher scores for these two dimensions compared to the experimental group. Furthermore, AI texts are more often 'penalised' in the control group, particularly among respondents with a humanities background and female respondents by receiving lower evaluations. However, the experimental group recorded higher creativity-related scores for AI texts in several cases, particularly for authors such as Mihai Eminescu and Lucian Blaga.

The stylistic and literary period characteristics of the Romanian authors included in the study are associated with differences in how poetry is received. It is worth noting that, while most participants were familiar with the selected authors due to their inclusion in the national curriculum, the specific poems used in the study were not necessarily well-known to them. Nevertheless, the literary prestige and recognisability of certain authors may have functioned as a confounding factor, potentially contributing to higher recorded evaluations of poems written by Romanian authors. Thus, adolescents seem to identify the texts of Mihai Eminescu and Lucian Blaga more easily, as they are recognisable by their 'classical' lyrical features. This correspondence is reflected in more critical evaluations of the AI-generated versions compared to the poetry of Nichita Stănescu and, in particular, Mircea Cărtărescu. In the case of Mircea Cărtărescu, the differences between the original text and the AI-generated version are recorded as less pronounced, given the colloquial and self-referential language, typical of postmodernist poetry. Although Mircea Cărtărescu's poetry receives predominantly positive but lower emotion-related evaluations among adolescents and is more difficult to differentiate from AI poetry, adolescents record higher creativity and complexity-related scores for it than AI poetry. This may be interpreted as reflecting a preference for a postmodern poetic language that is closer to the colloquial language used by young people and perceived as more accessible.

Educational profile is also associated with differences in poetic reception. Adolescents with a humanities background (philology or social sciences) record higher evaluations of human poetry. However, once they discover that a poem was generated by AI, their evaluations become more critical, as reflected in lower post-disclosure scores. During this re-evaluation, their level of disgust toward AI poetry is recorded as higher. However, respondents with a humanities background record higher evaluations of AI poetry when they are unaware of its authorship. By contrast, adolescents with a sciences background (mathematics-computer science, natural sciences) record comparatively less negative evaluations of AI poetry, although their emotion-related scores remain predominantly negative.

In general, gender differences are moderate but relevant in certain situations. Female respondents record higher evaluations of human poetry, particularly in terms of emotional reception, most clearly in the control group. By contrast, male respondents show fewer negative reactions toward AI texts in their evaluations. However, one of the most prevalent emotions associated in the evaluations with AI poetry is sadness. Perception of complexity and creativity are broadly similar between genders, with some variations depending on the author.

These findings highlight the role of labelling a poem as 'human' or 'AI-generated' in how it is evaluated. The analysis also shows that the reception of poetry by high school adolescents varies according to labels and templates; that is to say, the perception and evaluation of a poem vary according to prior information about its nature. In this sense, these variations may be understood as a locally form of 'patterned reception', reflecting pre-established interpretative frameworks shaped primarily by school-based literary learning practices rather than by text-internal features alone.

Although evaluations of poetry written by human authors in the control group were generally more favourable and AI poetry was initially better received in the experimental group, perceptions shifted negatively after the authorship was disclosed. This may reflect a broader cultural bias against the use of AI in literary creation, particularly in poetry, as well as to the role of author status and the expectations associated with it in the reception process. The emotional responses of adolescents in the experimental group to AI poetry are often recorded as more intense, which invites further reflection on the ability of these models to produce poetic content that blurs the boundaries between authenticity and originality in the age of algorithms. It remains to be seen how AI models will be integrated and accepted into the processes of poetic creation or co-creation, and above all, how future theories of reception will address this matter.

In the context of accelerating technological change, research into literary reception, and poetry in particular, is becoming increasingly relevant. This change may soon affect on a larger scale the way readers distinguish between literary creations by humans and those generated by AI. However, this debate could prevent us from reaching premature conclusions about the supposed end of human poetry. Understanding the effects of patterned reception and identifying textual triggers that are associated with differences in reader response and distinguish human poetry from AI-generated poetry are essential areas for reflection and analysis aimed at supporting the continuity of human poetry.

However, the limitations of this study must be acknowledged. Although the human-authored and AI-generated poems were carefully matched, it was not possible to make them perfectly equivalent. These differences may have affected how they were received by adolescents. In addition, as the data were collected from four intact classroom units, some of the observed patterns – particularly the consistency of evaluations and the synchronised shifts following authorship disclosure – can also be interpreted at the level of classroom contexts. Due to the small number of classes, no clustering adjustment was applied. The relatively small, localised sample of students also limits the scope of generalisation of the findings. Future studies could extend this research to larger and more diverse groups of participants, allowing for broader generalisation. In addition, inferential statistical analyses could offer deeper insights into the mechanisms underlying poetry reception and the impact of AI authorship on emotional and evaluative responses.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author/s.

Ethics statement

As the study involved human participants, it was conducted in accordance with established ethical research standards. Informed consent was obtained from all participants and/or their legal guardians prior to participation. Participation was entirely voluntary, and all responses were anonymized to ensure confidentiality. The study posed no psychological or physical risks to those involved.

Author contributions

AR: Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Amirjalili, F., Neysani, M., and Nikbakht, A. (2024). Exploring the boundaries of authorship: a comparative analysis of AI-generated text and human academic writing in English literature. *Front. Educ.* 9:1347421. doi: 10.3389/feduc.2024.1347421
- Barthes, R. (1967). *The death of the author*. London: Macat International Limited.
- Bellaiche, L., Shahi, R., Turpin, M. H., Ragnhildstveit, A., Sprockett, S., Barr, N., et al. (2023). Humans versus AI: whether and why we prefer human-created compared to AI-created artwork. *Cogn. Res. Princ. Implic.* 8:42. doi: 10.1186/s41235-023-00499-6
- Christou, P. (2023). A critical perspective over whether and how to acknowledge the use of artificial intelligence (AI) in qualitative studies. *Qual. Rep.* 28, 1981–1991. doi: 10.46743/2160-3715/2023.6407
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London: John Murray.
- Dobra, S., Halaszi, M., Kudor, D., and Medeșan, L. (2006). *Limba și Literatura Română. Manual pentru clasa a XI-a*. Bucharest: Corint Educational.
- Ebadi, S., Nejadghanbar, H., Salman, A. R., and Khosravi, H. (2025). Exploring the impact of generative AI on peer review: insights from journal reviewers. *J. Acad. Ethics* 23, 1383–1397. doi: 10.1007/s10805-025-09604-4
- Ekman, P., and Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs: Prentice-Hall.
- Grace, S. D., and Sahu, G. R. (2024). Navigating narrative frontiers: influence of generative AI on creative literature. *Int. Res. J. Adv. Eng. Manag.* 2, 1315–1323. doi: 10.47392/IRJAEM.2024.0179
- Hayles, N. K. (2012). *How we think: Digital media and contemporary Technogenesis*. Chicago: University of Chicago Press.
- Hitsuwari, J., Ueda, Y., Yun, W., and Nomura, M. (2023). Does human-AI collaboration lead to more creative art? Aesthetic evaluation of human-made and AI-generated haiku poetry. *Comput. Hum. Behav.* 139:107502. doi: 10.1016/j.chb.2022.107502
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Generative AI statement

The author(s) declared that Generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2026.1728281/full#supplementary-material>

Lakoff, G., and Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Limongi, R. (2024). The use of artificial intelligence in scientific research with integrity and ethics. *Future Studies Res. J.* 16:e845. doi: 10.24023/FutureJournal/2175-5825/2024.v16i1.845

Linardaki, C. (2022). Poetry at the first steps of artificial intelligence. *Humanist Studies Digital Age* 7:1–27. doi: 10.5399/uo/hstda/7.1.6

Nabasirye, D. (2024). The influence of author identity on literary reception. *European J. Literature Studies* 1, 13–23.

Porter, B., and Machery, E. (2024). AI-generated poetry is indistinguishable from human-written poetry and is rated more favorably. *Sci. Rep.* 14:26133. doi: 10.1038/s41598-024-76900-1

Saddhono, K., Judijanto, L., Nuryanto, T., Anis, M. Y., Nasution, J., and Sallu, S. (2024). "The AI-based design layout for exploring the reader experience through technical interaction for learning literature" in *In proceedings of the 4th international conference on advance computing and innovative Technologies in Engineering (ICACITE 2024)*, Greater Noida, Institute of Electrical and Electronics Engineers (IEEE). 1489–1494.

Selvi, V. T., and Ramya, P. (2024). Application of AI in literature: a study on evolution of stories and novels. *Recent Research Reviews J.* 3, 346–356.

Simion, E., Rogalski, F., and Cristea-Enache, D. (2007). *Limba și literatura română. Manual pentru clasa a XII-a*. Bucharest: Corint Educational.

Stockwell, P. (2002). *Cognitive poetics: A new introduction*. 1st Edn. London: Routledge.

Watson, S., and Romic, J. (2024). ChatGPT and the entangled evolution of society, education, and technology: a systems theory perspective. *Eur. Educ. Res. J.* 24, 205–224. doi: 10.1177/14749041231221266

Zwaan, R. A. (2004). "The immersed experienter: toward an embodied theory of language comprehension" in *The psychology of learning and motivation: Advances in research and theory*, ed. B. H. Ross, vol. 44 (Amsterdam: Elsevier Science), 35–62.