ALGORITHMIC NARRATIVES AND AFFECTIVE EXPRESSION: AN EXPLORATORY WITHIN-SUBJECT STUDY OF AI-ASSISTED WRITING AMONG STUDENTS AND PRE-SERVICE TEACHERS

Alfonso Filippone[1]
Department of Educational and Sports Sciences, Pegaso Telematic University
alfonso.filippone@unipegaso.it

Maria Ermelinda De Carlo
Department of Educational and Sports Sciences, Pegaso Telematic University
mariaermelinda.decarlo@unipegaso.it

Double Blind Peer Review

Citation
Filippone, A., & De Carlo, M. E. (2026). Algorithmic narratives and affective expression: An exploratory within-subject study of AI-assisted writing among students and pre-service teachers. *Italian Journal of Health Education, Sport and Inclusive Didactics, 10*(4)

Doi: https://doi.org/10.32043/gsd.v1i4.1906

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

gsdjournal.it
ISSN: 3103-5744

ABSTRACT
This study examines differences in emotional expression and narrative construction of the self between autonomous and AI-assisted writing in lower secondary school students and pre-service special education teachers. Framed within Digital Affective Literacy, the research adopted an exploratory mixed-methods within-subject design comparing narratives produced without AI with narratives generated through participant-created prompts. The sample included 50 students and 98 pre-service special education teachers. Data collection integrated the Tromsø Social Intelligence Scale, the Emotional Expressivity Scale and the Berkeley Expressivity Questionnaire. The analysis considered emotional richness, sentiment polarity, reflexivity, agency, relational complexity, perceived authenticity and narrative authorship. AI-assisted narratives were longer, more emotionally explicit and more structured, suggesting a tension between expressive amplification and possible narrative standardisation. These findings concern textual representation rather than direct evidence of deeper emotional processing or identity development.

KEYWORDS
Generative artificial intelligence; narrative self; emotional expression; digital affective literacy; narrative agency.

Received 2026-07-11
Accepted 2026-10-27
Published 2026-10-30

[1] Alfonso Filippone conceived the study, designed the protocol, conducted the data analysis, and drafted the manuscript. Maria Ermelinda De Carlo supervised the research and critically reviewed the final paper. Both authors approved the final version.

1. Introduction

The spread of generative artificial intelligence in educational contexts is transforming writing, learning, and meaning-making practices. Systems based on *Large Language Models* do not simply correct or complete texts, but intervene in the processes of ideation, reformulation, discursive organization, and linguistic representation of experience (Holmes et al., 2019; Luckin et al., 2016; Shi et al., 2025). In schools and training settings, these tools open up significant possibilities, but also raise pedagogical questions related to authorship, agency, authenticity, critical thinking, and the risk of cognitive delegation (Roe & Perkins, 2024; Park et al., 2025).

Recent literature on AI in education has focused primarily on personalization of learning, feedback, academic writing, language learning, integrity, plagiarism, and digital skills (Holmes et al., 2019; Li et al., 2025; Shi et al., 2025). The affective and narrative dimension of AI-mediated writing, however, appears less explored, especially when the task involves narrating one's self, emotions, and relationships. This gap is significant because narrative writing is not merely a linguistic exercise, but a practice of processing experience, constructing meaning, and organizing identity (Bruner, 1990; McAdams, 2001; Riessman, 2008).

From this perspective, generative AI can be interpreted not only as a technical tool to support writing, but as a potential symbolic mediator capable of influencing the way the subject represents themselves and their relationships. Through storytelling, in fact, the subject selects events, attributes meanings, organizes roles, constructs characters, and shapes emotional experiences (Bruner, 1990; McAdams, 2001). When this process is mediated by a generative

system, it becomes necessary to question what is expanded, reformulated, or standardized by the interaction with AI.

This study analyzed the impact of generative artificial intelligence on emotional expression and narrative construction of the self, comparing independently produced narratives with narratives created using AI based on participant-generated prompts. The sample included 50 students from two second-grade classes in middle school and 98 prospective specialization teachers enrolled in the Specialization Course for Teaching Support Activities in Middle School. The aim of the research was to understand whether and how the use of AI altered emotional richness, narrative structure, reflective depth, agency, and the perception of authenticity, also exploring the role of social intelligence and emotional expressiveness as individual variables associated with the observed transformations.

Accordingly, the study addressed the following research questions:

RQ1. How do autonomous and AI-assisted narratives differ in emotional richness, sentiment polarity, reflexivity, agency, relational complexity, and other textual characteristics?

RQ2. Do the differences between autonomous and AI-assisted narratives vary between lower secondary school students and pre-service special education teachers?

RQ3. How are individual differences in social intelligence and emotional expressiveness associated with the observed differences between the two writing conditions?

RQ4. How do participants perceive the authenticity, emotional involvement, and authorship of AI-assisted narratives, and what recurrent patterns of

expressive amplification or narrative standardisation emerge from the comparison of the texts?

2. Theoretical framework

2.1 Generative AI, writing and agency

Artificial intelligence in education was initially studied in relation to personalization of learning, intelligent tutors, adaptive systems, and feedback automation (Luckin et al., 2016; Holmes et al., 2019). With the advent of generative AI, the relationship between technology and learning has taken on a new configuration: conversational systems produce texts, explanations, dialogues, and narratives starting from instructions formulated in natural language (Shi et al., 2025). The most recent reviews show that LLMs are used to support writing, feedback, language learning, tutoring, and instructional design, but also highlight critical issues related to reliability, tool dependency, integrity, and critical use (Li et al., 2025; Law, 2024; Shi et al., 2025)

In writing, these tools can support ideation, organization, expansion, revision, and reformulation. They can broaden the vocabulary, make the text structure more coherent, and reduce initial difficulties in writing (Li et al., 2025; Law, 2024). However, the use of AI raises crucial questions about agency: if a significant portion of the text is generated by a system, we must ask who retains control over the meaning and to what extent the final product is recognized as the subject's own (Roe & Perkins, 2024).

In the case of emotional writing, this question takes on specific relevance. The text is not about neutral content, but about a relational and affective experience. AI can help name difficult emotions or make a narrative more complex, but it can also produce generic expressive formulas, attenuate the ambivalence of experience, or replace the personal voice with a linguistically

more fluid but less authentic form. For this reason, reflection on AI in writing must include not only the quality of the product, but also agency, intentionality, and narrative authorship (Bruner, 1990; McAdams, 2001; Roe & Perkins, 2024).

2.2 Emotion, learning and self-narration

The emotional dimension is an integral part of learning. Emotions guide attention, memory, situational evaluation, and the attribution of meaning; therefore, learning, cognition, and affective experience cannot be separated (Immordino-Yang & Damasio, 2007). This perspective is consistent with Social and Emotional Learning models, which emphasize the importance of self-awareness, self-regulation, social awareness, relational skills, and responsible decision-making (CASEL, 2020).

Narration constitutes a privileged space for processing this complexity. According to Bruner (1990), narrative thinking allows individuals to organize experience through intentions, actions, contexts and meanings. McAdams (2001) highlighted the role of life stories in the construction of personal identity, while Riessman (2008) showed how autobiographical narratives can be analyzed as social and interpretative practices through which the subject shapes the self.

In this study, the request to produce a story, fable, or fairy tale about their relationship with relationships allowed participants to express personal experiences in symbolic or metaphorical form. This methodological choice is based on the idea that narrative allows for the representation of emotions, conflicts, and relationships in an indirect yet meaningful way (Bruner, 1990; McAdams, 2001; Riessman, 2008). The comparison with AI-mediated production allowed us to observe how the narrative construction of the self-

changes when the subject formulates a prompt through which he entrusts the generative system with the reworking of his experience.

2.3 Digital Affective Literacy

The increasing use of generative AI in emotionally significant writing requires an extension of current approaches to digital and AI literacy. Digital literacy broadly concerns the competent, critical, and responsible use of digital technologies, while AI literacy focuses more specifically on understanding how artificial intelligence systems operate, critically evaluating their outputs, and interacting with them in informed and ethical ways (Long & Magerko, 2020; Chiu et al., 2024). Neither construct, however, focuses specifically on how generative systems mediate the linguistic representation of emotions, relationships, and personal experience.

Digital Affective Literacy (DAL) is proposed in this study as a pedagogical and analytical framework situated at the intersection of AI literacy, emotional literacy, social-emotional learning, and narrative competence. Emotional literacy and social-emotional learning concern the ability to recognise, understand, express, and regulate emotions in personal and interpersonal contexts (CASEL, 2020). Narrative competence concerns the ability to organise experience through characters, intentions, events, temporal connections, and meanings (Bruner, 1990; McAdams, 2001). DAL specifically addresses what happens when these affective and narrative processes are mediated by a generative system.

DAL is therefore defined as the ability to recognise, interpret, and critically govern the ways in which generative technologies transform the linguistic representation of emotions, relationships, and personal experience. It includes the capacity to examine what the system adds, removes, intensifies,

attenuates, reorganises, or standardises; to evaluate whether the generated language corresponds to the author's intended meaning; and to maintain meaningful control over the resulting narrative. This focus is particularly relevant because AI-mediated communication can reshape how individuals formulate and present their experiences, while different levels of human contribution may also influence perceived authorship and psychological ownership (Hancock et al., 2020; Draxler et al., 2024).

The framework comprises four interconnected dimensions: affective awareness, referring to the recognition of emotions and relational meanings; representational awareness, concerning the transformation of experience into language; algorithmic awareness, involving an understanding of how generative systems shape textual outputs; and authorial agency, referring to the capacity to accept, modify, reject, or critically appropriate AI-generated formulations. DAL is not presented here as an already validated psychological construct or measurement scale. It is proposed as an emerging conceptual framework for analysing and teaching critical awareness of AI-mediated affective and narrative writing. In educational contexts, its purpose is to support the use of generative AI without obscuring the learner's voice, interpretative responsibility, or control over meaning.

2.4 Emotional amplification and the risk of standardization

One of the study's central questions concerns the dual effect of generative AI on emotional writing. On the one hand, AI can function as an expressive amplifier: it suggests words for difficult-to-name emotions, clarifies the relationships between events and internal states, and helps organize the narrative into a coherent form (Law, 2024; Li et al., 2025). On the other hand, it can introduce forms of standardization, producing fluid, balanced texts oriented toward reassuring narrative closures, but sometimes less situated and

less aligned with the ambivalence of personal experience (Roe & Perkins, 2024; Shi et al., 2025).

For this reason, the study did not limit itself to verifying whether the texts generated with AI were longer or lexically richer, but also analysed reflexivity, agency, relational complexity and perception of authenticity, to understand whether AI supported the construction of meaning or tended to replace the narrative voice of the subject (Bruner, 1990; McAdams, 2001; Roe & Perkins, 2024).

2.5 Individual differences: social intelligence and emotional expression

The effect of AI on narrative writing cannot be considered uniform. It is plausible that it varies in relation to some individual characteristics, particularly social intelligence and emotional expressiveness. Social intelligence concerns the ability to understand social situations, predict others' reactions, and interpret relational dynamics; in the Tromsø Social Intelligence Scale, it is divided into social information processing, social skills, and social awareness (Silvera et al., 2001).

Since the proposed narrative task concerned relationships, this variable could impact the quality of the prompts, the complexity of the characters, and the representation of interpersonal dynamics. A participant with greater social reading skills could, in fact, provide the AI with more detailed instructions and produce, or have it produce, more relationally complex narratives (Silvera et al., 2001; Riessman, 2008).

Emotional expressivity was considered through two complementary instruments: the *Emotional Expressivity Scale* (EES), which measures the global tendency to externally display emotions (Kring et al., 1994), and the *Berkeley Expressivity Questionnaire* (BEQ), which distinguishes positive expressivity,

negative expressivity, and strength of emotional impulse (Gross & John, 1997). The inclusion of these measures allowed us to explore whether dispositional emotional expressiveness was associated with the magnitude of the textual differences observed between autonomous and AI-assisted narratives.

In the present study, emotional expressiveness, emotional richness, and emotional density were treated as distinct constructs. Emotional expressiveness referred to the participant's dispositional tendency to externally display emotions and was assessed through the EES and BEQ. Emotional richness was considered a textual characteristic and referred to the range and variety of emotional meanings represented in a narrative, as reflected in the number and diversity of emotional markers. Emotional density referred to the relative concentration of explicit emotional markers in relation to textual length and was calculated per 100 words. Emotional expressiveness therefore represented an individual characteristic, whereas emotional richness and emotional density described characteristics of the narratives.

2.6 Research hypotheses

Building on the theoretical framework and the research questions presented above, the quantitative component of the study was guided by the following hypotheses:

H1. AI-assisted narratives will exhibit greater emotional richness and emotional density than autonomous narratives.

H2. AI-assisted narratives will be longer and more structurally regular than autonomous narratives, while also displaying recurrent linguistic or thematic patterns potentially consistent with narrative standardisation.

H3. AI-assisted narratives will display higher textual indicators of reflexivity, agency, and relational complexity than autonomous narratives.

H4. The magnitude of the differences between autonomous and AI-assisted narratives will be associated with individual differences in social intelligence and emotional expressiveness.

3. Material and methods

3.1 Research design

The study adopted an exploratory mixed-methods within-subject design aimed at comparing the emotional and narrative characteristics of autonomous and AI-assisted writing. Each participant produced two narratives addressing the same relational theme. The first narrative was written autonomously without AI support. The second was produced with the support of a generative AI system through one or more prompts formulated by the participant. The within-subject comparison made it possible to examine differences between the two writing conditions within the same participant, thereby reducing the influence of stable inter-individual variability.

The quantitative component compared textual and narrative indicators across the two conditions and explored their associations with social intelligence and dispositional emotional expressiveness. The qualitative component examined recurrent patterns in the paired narratives and participants' metacognitive reflections, with particular attention to expressive amplification, narrative organisation, perceived authenticity, and authorship. Quantitative and qualitative findings were integrated during the interpretation of the results.

All participants completed the autonomous narrative before the AI-assisted narrative. The order of the two conditions was therefore fixed rather than randomised or counterbalanced. Consequently, the observed differences cannot be attributed exclusively to AI assistance, as practice, familiarity with the task, further reflection on the relational experience, and the opportunity to

reformulate the previous narrative may also have contributed. The results are therefore interpreted as differences associated with the AI-assisted writing condition rather than as exclusively causal effects of generative AI.

The overall methodological design is summarised in Figure 1.

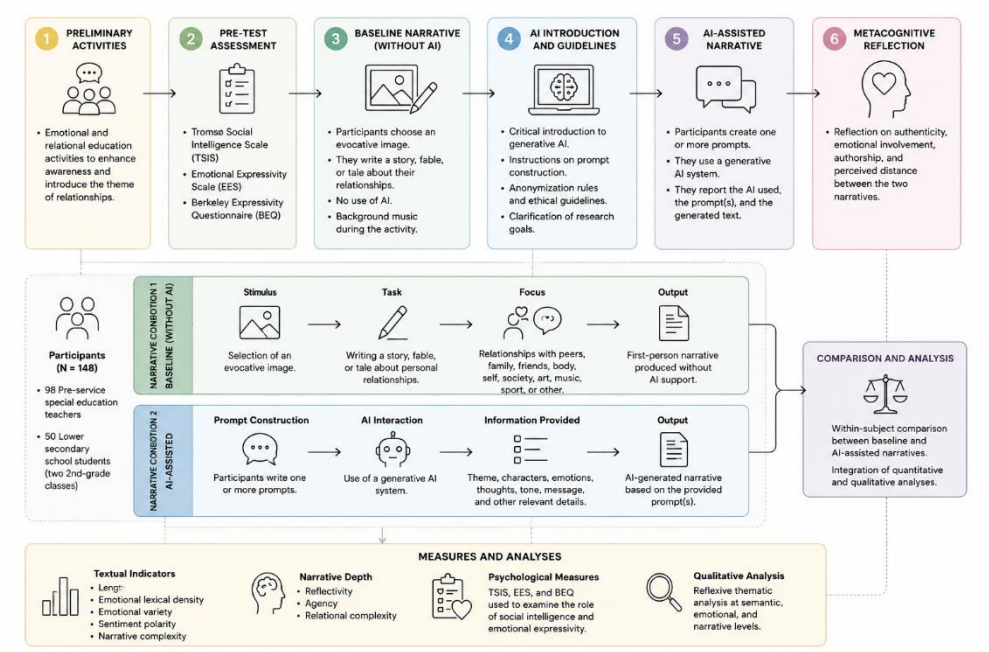


Figure 1: Overview of the mixed-methods within-subject design and sequence of the study.

3.2 Participants

The sample comprised 148 participants divided into two educational groups. The first group included 98 pre-service special education teachers enrolled in the Specialisation Course for Teaching Support Activities in Lower Secondary Education. The second group included 50 lower secondary school students attending two second-year classes.

Participants were selected through purposive sampling on the basis of their involvement in the educational activities within which the study was conducted. The final analytic sample included participants whose autonomous and AI-assisted narratives, questionnaire responses, and metacognitive reflections could be matched through the same anonymous identification code.

The inclusion of the two groups allowed the study to explore AI-mediated emotional writing in two different educational contexts: adolescents engaged in processes of personal and relational development, and adults preparing for a professional role involving reflective, inclusive, and socio-relational competencies. The comparison was exploratory and was not intended to produce population-level estimates or to imply that the two groups were developmentally equivalent.

3.3 Instruments and narrative tasks

Social Social intelligence was assessed using the Tromsø Social Intelligence Scale (TSIS; Silvera et al., 2001). The TSIS comprises 21 items rated on a seven-point response scale and is organised into three dimensions: social information processing, social skills, and social awareness. In the present study, it was used to examine individual differences in the understanding of social situations and interpersonal dynamics.

Dispositional emotional expressiveness was assessed through two complementary instruments. The Emotional Expressivity Scale (EES; Kring et al., 1994) comprises 17 items and provides a global measure of the tendency to externally express emotions. The Berkeley Expressivity Questionnaire (BEQ; Gross & John, 1997) comprises 16 items and distinguishes positive expressivity, negative expressivity, and impulse strength. The EES and BEQ therefore

assessed participant-level characteristics rather than textual properties of the narratives.

Narrative data were collected through two writing tasks concerning a personally meaningful relational experience. In the autonomous condition, participants selected an evocative image and wrote a story, fable, or fairy tale without using artificial intelligence. The relational domain could concern peers, family, friendships or affective relationships, the social context, the body, the self, art, music, sport, or another personally significant relationship. A background melody accompanied the writing activity to support concentration and introspection.

The autonomous writing instruction was formulated as follows: “Starting from the selected image, write a story, fable, or fairy tale concerning a relationship that is meaningful to you. You may refer to your relationship with other people, with yourself, with your body, or with another personally significant area of experience. Describe the event in a way that represents your perspective, without using artificial intelligence.”

In the AI-assisted condition, participants used ChatGPT to produce a new version of the narrative through one or more prompts formulated by them. They were instructed to include information about the relational theme, characters, emotions, thoughts, narrative tone, and intended final message, while avoiding personally identifying information. Participants recorded the prompts used, the generated narrative, the number of prompts, and any personal modifications made to the generated text. The specific ChatGPT model version was not recorded and is therefore considered a methodological limitation. After completing the AI-assisted task, participants provided a brief open-ended metacognitive reflection. The reflection explored the extent to which the generated narrative represented their intended experience, the

changes introduced by AI, their emotional involvement, their perception of authenticity, and their sense of narrative authorship. These responses were used as qualitative material and were not treated as scores derived from a validated authenticity or authorship scale.

3.4 Procedure

The procedure was organised into six sequential phases and was implemented in the same order for all participants. First, participants took part in preliminary educational activities concerning emotions and relationships. These activities introduced the topic of relational experience and supported the recognition and verbalisation of emotional states.

Second, participants completed the TSIS, EES, and BEQ. Each questionnaire and narrative record was associated with the same anonymous identification code, allowing the materials produced in the different phases to be matched without using personally identifying information.

Third, participants completed the autonomous narrative task. After selecting an evocative image, they wrote a story, fable, or fairy tale concerning a personally meaningful relational experience. They were explicitly instructed not to use generative AI during this phase. A background melody accompanied the activity to support concentration and introspection.

Fourth, participants received a critical introduction to generative AI. The activity addressed the basic functioning and limitations of generative systems, the construction of prompts, the protection of personal information, and the distinction between using AI as a writing support and delegating authorship to the system. Participants were instructed not to enter real names, recognisable places, or other identifying information.

Fifth, participants completed the AI-assisted narrative task using ChatGPT. They formulated one or more prompts that included the relational theme, characters, emotions, thoughts, narrative tone, and intended final message. They recorded the prompts used and the generated text. After reading the output, they could retain or personally modify it and were asked to report any changes made. The number of prompts and the declared personal modifications were retained as process-related information.

Sixth, participants completed an open-ended metacognitive reflection concerning the differences between the two narratives, the extent to which the AI-assisted text represented their intended experience, their emotional involvement, their perception of authenticity, and their sense of authorship. Because the same sequence was followed by all participants, the autonomous condition always preceded the AI-assisted condition.

3.5 Data analysis

The analysis combined quantitative textual indicators, psychometric measures, and qualitative interpretation. The autonomous and AI-assisted narratives were linked at the participant level through the anonymous identification code and analysed as paired texts.

Textual length was measured through the number of words in each narrative. Emotional richness was operationalised through the frequency and variety of explicit emotional markers. Emotional density was calculated as the number of emotional markers per 100 words, thereby accounting for differences in textual length. Emotional variety referred to the number of different emotional categories represented in the narrative. Sentiment polarity was used as a descriptive indicator of the relative presence of positively and negatively oriented emotional language.

Narrative depth was operationalised through reflexivity, agency, and relational complexity. Reflexivity referred to textual passages in which participants connected events, thoughts, emotions, and personal meanings, including expressions of awareness, interpretation, or reappraisal. Agency referred to passages representing the protagonist as making choices, initiating actions, confronting difficulties, or contributing to change. Relational complexity concerned the representation of interpersonal perspectives, intentions, reciprocal responses, conflicts, ambivalence, and relational transformations. These indicators described characteristics of the texts and were not treated as direct measures of participants' psychological functioning.

A structured coding framework was used to identify the narrative indicators. The coding unit could consist of a word, phrase, or clause, depending on the indicator. Expressions such as "I realised" or "I understood why" were treated as potential reflexive markers when they explicitly connected experience and meaning. Expressions such as "I decided," "I chose," or "I tried" were considered potential agency markers when they represented intentional action. References to another person's perspective, emotional response, or change in the relationship were considered relational markers. Coding decisions were based on their meaning within the narrative context rather than on the isolated presence of individual words. These definitions and contextual examples constituted the coding framework applied consistently to the autonomous and AI-assisted narratives. Narrative standardisation was examined qualitatively through the recurrence of similar structures, formulations, and reassuring or growth-oriented narrative conclusions, and was not treated as a validated quantitative measure.

Scale scores were calculated according to the scoring procedures of the respective instruments, including the recoding of reverse-worded items where

required. Internal consistency was examined through Cronbach's alpha for the total scales and relevant subdimensions. Descriptive statistics were calculated for the overall sample and separately for students and pre-service special education teachers.

Differences between autonomous and AI-assisted narratives were examined through paired-samples t-tests. When the assumptions required for parametric analysis were not met, the corresponding non-parametric procedure was considered. Cohen's dz was used to describe the magnitude of within-participant differences. Change scores were calculated by subtracting the autonomous-condition value from the AI-assisted-condition value. Differences in change scores between students and pre-service teachers were examined using independent-group comparisons, with Cohen's d used as the effect-size measure.

Associations between psychometric characteristics, prompt-related variables, and changes in narrative indicators were examined using Spearman's rank correlations. The analyses were exploratory and were interpreted by considering the direction and magnitude of the associations together with the corresponding p values and effect sizes. No formal adjustment for multiple comparisons was applied; consequently, individual significance tests were not treated as independent confirmatory evidence. References to regression models were removed because the reported findings concerned paired comparisons, group comparisons, and correlational analyses.

The qualitative component followed the principles of reflexive thematic analysis (Braun & Clarke, 2006, 2019, 2021). The analysis involved familiarisation with the paired narratives and metacognitive reflections, initial coding, development of candidate themes, review of the themes across cases and groups, definition and naming of the themes, and interpretative

integration. The analysis considered three interconnected levels: the semantic level, concerning the relational contexts and meanings explicitly represented; the emotional level, concerning the quality, variety, and articulation of emotions; and the narrative level, concerning structure, agency, relational dynamics, and narrative closure.

The first author conducted the initial familiarisation, coding, and theme development. The second author critically reviewed the emerging thematic structure and its coherence with the narrative material. Interpretative differences were discussed reflexively and used to refine the definitions and boundaries of the themes. This process was intended to strengthen transparency and interpretative coherence rather than to treat qualitative coding as an objectively fixed classification. Sentiment polarity was used as a descriptive support and not as a substitute for qualitative interpretation. Quantitative and qualitative findings were integrated through comparison of convergent, complementary, and contrasting patterns across the two writing conditions.

3.6 Ethical aspects

The research was conducted in accordance with the ethical principles of educational research. Under the applicable institutional regulations, formal ethics committee approval was not required because the study was conducted as a non-clinical educational activity involving voluntary participation and anonymised data. Participation was voluntary and based on informed consent. Adult participants provided their consent directly, while written authorisation was obtained from the parents or legal guardians of the participating students.

Participants were asked not to include real names, recognizable places, or sensitive information about themselves or others in the texts and prompts. The

activity was presented as a reflection and research experience, not as an assessment of the personal quality of their writing. It was also made clear that there were no right or wrong answers and that the purpose of the study was to compare emotional and narrative representation in autonomous and AI-assisted writing.

4. Results and discussion

4.1 Preliminary data checks and internal consistency of the instruments

The overall sample consisted of 148 participants: 50 students from two second-year classes in middle school and 98 prospective support teachers enrolled in the Specialization Course for Teaching Support Activities in Middle School. All participants were present in the datasets for baseline narration, narration with AI, TSIS, BEQ, and EES. No missing or duplicate IDs were detected.

In the AI condition, all participants used ChatGPT. The number of prompts used ranged from two to four, with a prevalence of three prompts. The most frequently represented relational areas concerned relationships with peers, self-esteem, family, and, in the group of future support teachers, relationships with work, educational roles, and professional dynamics.

Before calculating the scores, the items worded in the opposite direction were recoded, so that higher values indicated higher levels of the measured constructs. In particular, the TSIS showed adequate to excellent internal consistency in the overall sample and across its three subdimensions (Silvera et al., 2001). The EES and BEQ likewise showed adequate to excellent internal consistency in the present sample (Kring et al., 1994; Gross and John, 1997).

Table 1: *Internal reliability of the psychometric instruments used in the study.*

Tool	Dimension	Cronbach's α
TSIS	Total	0.91
TSIS	Information processing	0.76
TSIS	Social skills	0.79
TSIS	Social awareness	0.76
EES	Global emotional expressiveness	0.94
BEQ	Total	0.91
BEQ	Positive expressiveness	0.84
BEQ	Negative expressiveness	0.87
BEQ	Strength of the emotional impulse	0.86

4.2 Textual and emotional changes between baseline and AI

The comparison between the two conditions highlighted a significant increase in the length of the texts produced with AI. As summarized in Table 2, the narratives produced with AI showed higher average values in all the textual, emotional, and narrative indicators considered.

Table 2: Comparison between baseline and AI-assisted narratives in the overall sample

Indicator	Baseline M	AI M	Test	*p*	Cohen's dz
Number of words	70.32	125.09	$t(147) = 45.34$	< 0.001	3.73
Emotional markers	0.59	1.66	$t(147) = 19.48$	< 0.001	1.60
Emotional density	0.88	1.32	$t(147) = 7.58$	< 0.001	0.62
Reflexivity	6.78	9.42	$t(147) = 11.86$	< 0.001	0.97
Agency	5.05	9.71	$t(147) = 26.12$	< 0.001	2.15
Relational markers	1.16	1.93	$t(147) = 14.45$	< 0.001	1.19

All six paired comparisons were statistically significant. The largest standardized change concerned text length (dz = 3.73), with AI-assisted narratives being substantially longer than baseline narratives. Higher values were also observed for emotional markers, emotional density, reflexivity,

agency, and relational markers. These differences should be interpreted as changes in observable textual and narrative features rather than as direct evidence of changes in participants' underlying emotional or psychological states. Accordingly, the increases in emotional markers and emotional density are consistent with H1 at the level of linguistic expression, indicating greater explicit verbalization of affective content in the AI-assisted condition.

This finding is consistent with the literature on generative AI in writing, which highlights the ability of linguistic systems to support ideation, reformulation, lexical expansion, and textual organization (Law, 2024; Cheng et al., 2025; Li et al., 2025; Shi et al., 2025). In the case of emotional writing, this support takes on a specific pedagogical significance, as it affects the way in which the subject names, organizes, and communicates his or her emotional experience.

4.3 Reflexivity, agency and relational complexity

The data relating to reflexivity, agency, and relational markers, reported in Table 2, show that the narratives produced with AI were not only longer, but also more explicit in representing reflective passages, the protagonist's actions, and interpersonal dynamics. In particular, the AI-assisted texts tended to clarify the relationships between emotions, thoughts, behaviors, and final transformations.

From a theoretical perspective, this result can be interpreted in light of the relationship between narration and the construction of the self. According to Bruner (1990), storytelling allows the subject to organize experience through intentions, events, and meanings. McAdams (2001) has shown how personal stories contribute to the construction of identity, while Riessman (2008) has emphasized the interpretative and social nature of narration. From this perspective, the textual patterns observed in the AI-assisted condition may

reflect a more explicit narrative organization of experience, particularly in the representation of emotions, relationships, and connections between events.

However, this very increased organization raises a critical issue. Many AI-assisted narratives presented similar interpretative closures, oriented toward personal growth, understanding others, reconciliation, or the positive transformation of conflict. This suggests a possible standardization effect: AI tends to make the text more fluid and coherent, but can also reduce complex experiences to more predictable narrative forms. In this sense, narrative improvement at the formal level does not necessarily correspond to a deeper or more authentic elaboration of lived experience. The educational challenge, therefore, is to help learners critically compare the AI-generated version with their own initial narrative voice. This aspect recalls the debate on agency in AI-mediated writing and the need to preserve the subject's authorial control over the meaning produced (Roe & Perkins, 2024).

4.4 Thematic analysis of narratives

The reflexive thematic analysis, conducted according to Braun and Clarke (2006), identified recurring interpretative patterns across the two participant groups and the two writing conditions, including self-recognition, the need to belong, fear of judgment, relational conflict, the search for trust, the desire to be understood, and the transformation of negative experiences into opportunities for growth. These themes were developed through iterative comparison of the coded extracts, the complete narratives, and the paired baseline and AI-assisted texts. Figure 2 provides an interpretative synthesis of the main cross-condition transformations identified through this analytic process.

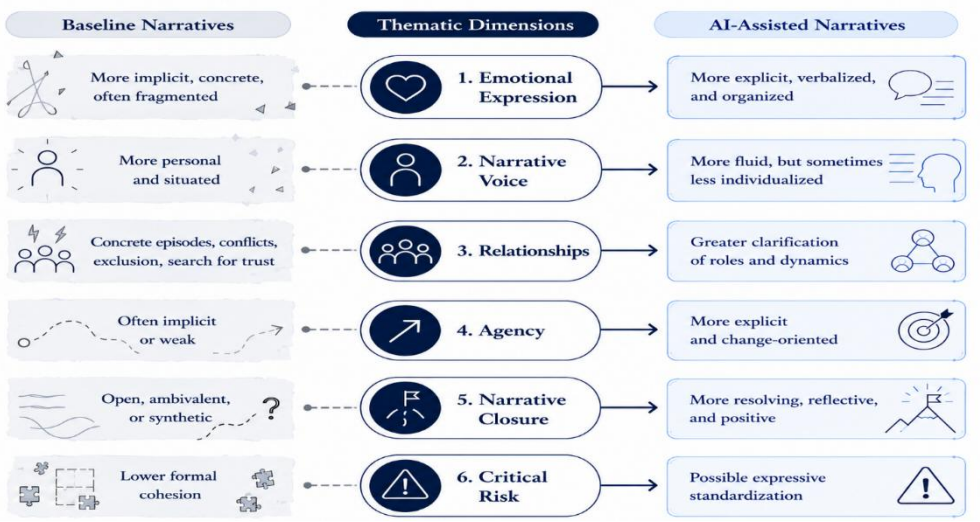


Figure 2: Interpretative synthesis of the main thematic transformations identified through reflexive thematic analysis across baseline and AI-assisted narratives.

As summarized in Figure 2, the AI-assisted condition was characterized not only by quantitative textual expansion but also by recurrent qualitative differences in narrative form: experiences were generally represented more explicitly and sequentially and were more frequently oriented toward reflective closure, although some texts were interpreted as less singular in voice than their baseline counterparts.

In the student group, baseline narratives often focused on peer relationships, friendship, exclusion, shame, and the desire for acceptance. The texts were shorter and sometimes fragmented, but retained a strong expressive immediacy: emotions emerged through concrete images, symbolic situations, or characters' actions rather than through direct verbalization. In the AI-assisted narratives, the same core narratives were reorganized into more explicit and coherent forms: exclusion often became a path to personal recognition, conflict an opportunity for dialogue, solitude a chance to discover one's own value. This pattern is consistent with greater emotional verbalization

in the AI-assisted narratives, while also revealing a recurrent tendency toward more positively resolved narrative trajectories.

In the group of future support teachers, the baseline narratives presented themes related to care, responsibility, inclusion, the educational relationship, and the professional role. The relationship with the self was often intertwined with the relationship with the vulnerable other, with the class group, with the disability, or with the strain of the helping relationship. In the AI condition, these cores were more organized and oriented toward the construction of an educational message, with recurring key words such as listening, empathy, inclusion, resilience, and care. Even in this group, however, a risk of uniformity emerged, as the complexity of the educational relationship sometimes tended to be reduced to idealized formulas.

Overall, the comparison between baseline and AI highlights a transformation from situated narration to explicit narration. In the baseline, concreteness, fragmentation, personal images, and ambivalence prevailed; in the AI condition, order, cohesion, emotional explicitness, and reflective closure prevailed. This transformation should not be read in simply improving terms: AI-assisted narratives were generally more explicit and structured, while some were interpreted as less singular in voice and less marked by the ambiguity observed in the corresponding baseline narratives. In line with Bruner (1990), McAdams (2001), and Riessman (2008), the value of narration does not coincide with the formal success of the text, but with the possibility of recognizing oneself in the form given to the experience.

4.5 Differences between students and future support teachers

The pattern of change between baseline and the AI-assisted condition differed between the two participant groups. Both groups produced longer texts in the

AI-assisted condition, while the increase in emotional density was more pronounced among students than among future support teachers. Among students, emotional density increased from 1.58 to 2.26, corresponding to a mean change of 0.679, whereas among future support teachers it increased from 0.52 to 0.84, corresponding to a mean change of 0.314. The between-group difference in change in emotional density was statistically significant, Welch's $t(79.10) = 2.83$, $p = 0.006$, Cohen's $d = 0.54$, mean difference = 0.365, 95% CI [0.108, 0.622].

Students also showed larger increases in the coded indicators of reflexivity, agency, and relational references. At the textual level, this pattern suggests that the difference between baseline and AI-assisted narratives was more pronounced among adolescents, whose autonomous narratives tended to be more concise and implicit. Among future support teachers, baseline narratives already contained comparatively more reflective and relational elaboration, and the observed increase in the AI-assisted condition was therefore less marked.

One possible interpretation of this group difference concerns the participants' different developmental and formative contexts. Middle school students are in a phase in which identity construction and the reworking of relationships with peers, family, and self-take on particular importance. In this phase, narrative can foster the symbolization of emotional and relational experiences. Future support teachers, on the other hand, are adults in professional training, more accustomed to reflecting on educational and relational dynamics. Accordingly, the observed textual pattern may be described as more compensatory among students and more reorganizational among future support teachers, although this interpretation remains exploratory.

This finding is consistent with the Social and Emotional Learning perspective, according to which emotional awareness, social awareness, and relational skills are developed through intentional and reflective educational practices (CASEL, 2020). In this regard, AI can represent a teaching resource only if accompanied by pedagogical mediation, capable of transforming textual production into an opportunity for emotional awareness and not simply the automatic generation of content.

4.6 Emotional expressiveness and narrative transformations

Exploratory individual-differences analyses identified several associations between dispositional emotional expressiveness and changes in narrative indicators from baseline to the AI-assisted condition. Higher overall BEQ scores were associated with smaller increases in emotional density (Spearman's $\rho = -0.246$, $p = 0.003$). Similar negative associations were observed for positive expressiveness ($\rho = -0.216$, $p = 0.008$) and negative expressiveness ($\rho = -0.271$, $p < 0.001$). Negative expressiveness was also negatively associated with change in reflexivity markers ($\rho = -0.234$, $p = 0.004$).

These associations indicate that participants with higher dispositional expressiveness scores tended to show smaller increases in some of the coded textual indicators between the two writing conditions. Conversely, larger changes in emotional density were more frequently observed among participants with comparatively lower BEQ scores. This pattern may be consistent with differential changes across participants, but it should not be interpreted as evidence that AI exerted a compensatory psychological effect or that participants with lower expressiveness necessarily benefited more from AI.

An additional exploratory association concerned interaction with the AI system. The number of prompts used was positively associated with the increase in text length ($\rho = 0.246$, $p = 0.003$), indicating that participants who engaged in more prompting tended to show larger differences in word count between baseline and AI-assisted narratives. This association does not establish a causal effect of prompting, since prompt number may also reflect participants' writing strategies, engagement with the task, or characteristics of the narrative being developed.

Overall, these findings suggest that the magnitude of textual change observed in the AI-assisted condition was not uniform across participants and was associated with selected dispositional and interaction-related variables. Given the exploratory nature of these analyses and the number of associations examined, they should be interpreted cautiously and as indications for further investigation rather than as evidence of underlying psychological mechanisms.

4.7 Authenticity, authorship and narrative standardization

Authenticity was not treated in this study as an objectively measurable property of the narratives. Rather, it was approached at two complementary levels. First, participants' metacognitive reflections were used to explore their subjective perceptions of authenticity, emotional involvement, narrative authorship, and perceived distance between the autonomous and AI-assisted versions. Second, the researchers examined differences between paired narratives in terms of voice, structural organization, emotional explicitness, and recurrent narrative trajectories. These two levels were kept analytically distinct: participants' perceptions concerned their subjective relationship with the text, whereas observations regarding standardization and narrative form derived from the researchers' qualitative interpretation of the corpus.

At the textual level, AI-assisted narratives were generally more explicit, orderly, and narratively cohesive. The comparative analysis also identified recurrent transformations, including greater explicitation of emotions, clearer articulation of interpersonal relationships, and more frequent use of concluding reflections. In several narratives, conflicts were reorganized into trajectories oriented toward personal growth, understanding, reconciliation, or positive reappraisal. These recurring patterns were interpreted as possible indicators of structural homogenization rather than as direct evidence of reduced authenticity.

Accordingly, greater formal coherence should not be equated with greater narrative authenticity. A narrative may become more fluent and communicatively effective while simultaneously departing, to varying degrees, from the participant's initial linguistic or narrative style. Conversely, AI-mediated reformulation may also enable participants to articulate meanings or emotions that were only implicit in the autonomous text. Authorship should therefore be understood here as a relational and processual dimension concerning the extent to which participants continue to recognize themselves in, intentionally shape, and critically evaluate the AI-assisted narrative.

These findings suggest a tension rather than a simple opposition between expressive support and authorial displacement. Generative AI may facilitate emotional verbalization and narrative organization, while also introducing recurrent linguistic or interpretative patterns that risk reducing singularity and ambivalence. For educational purposes, the relevant question is therefore not whether an AI-assisted text is intrinsically authentic or inauthentic, but whether learners are able to identify what has been retained, reformulated, added, or displaced through algorithmic mediation and to maintain critical ownership of the resulting narrative.

This perspective is consistent with the need for a form of critical and affective AI literacy in which learners are supported not only in producing texts with generative systems, but also in interrogating the relationship between personal experience, linguistic representation, and algorithmic mediation. Within the framework proposed in this study, this reflective capacity constitutes a central component of Digital Affective Literacy (Cheng, et al., 2025; Holmes et al., 2019; Park et al., 2025; Shi et al., 2025).

4.8 Summary of hypotheses and pedagogical implications

Overall, the findings provide differentiated support for the study hypotheses. H1 was supported at the level of observable textual indicators, as AI-assisted narratives showed higher emotional markers and emotional density than baseline narratives; these differences indicate greater linguistic explicitness of affective content rather than greater experienced emotional intensity. H2 received partial support: AI-assisted narratives were generally longer and more formally structured, while the qualitative analysis also identified recurrent narrative trajectories consistent with possible linguistic and thematic standardization. H3 was supported at the textual level by higher coded values for reflexivity, agency, and relational markers in the AI-assisted condition; these changes should not be interpreted as direct evidence of corresponding psychological changes. H4 received partial and exploratory support: selected dimensions of dispositional emotional expressiveness measured by the BEQ were associated with changes in emotional density and reflexivity, whereas the available analyses did not provide comparable evidence for social intelligence. H5 received exploratory qualitative support. Participants' metacognitive reflections addressed perceived authenticity, emotional involvement, and narrative authorship, while the researchers' comparative analysis identified a tension between greater expressive organization and recurrent patterns of

narrative standardization. These findings concern perceived and interpreted dimensions of authorship and authenticity rather than objective measures of either construct.

The pedagogical implications therefore require a balanced interpretation. AI-assisted narrative writing may provide useful support for emotional verbalization, textual organization, and reflective comparison between alternative ways of representing experience. At the same time, greater reliance on algorithmic mediation may introduce risks of narrative homogenization, attenuation of ambiguity, and reduced authorial control when generated formulations are accepted without critical examination. The educational value of generative AI should therefore not be identified simply with the production of longer, more explicit, or formally coherent texts. Rather, its pedagogical potential may lie in using the comparison between autonomous and AI-assisted writing as an opportunity to examine what the system adds, reformulates, simplifies, or displaces. From this perspective, Digital Affective Literacy (DAL) involves not only the ability to use generative technologies, but also the capacity to critically evaluate their mediation of emotional language, preserve narrative agency, and maintain reflective ownership of the meanings expressed.

5. Conclusions

This study examined differences between autonomously produced narratives and AI-assisted narratives in emotional and relational writing among middle school students and future support teachers. Across the two conditions, AI-assisted narratives were generally longer, more emotionally explicit, and characterized by higher coded values for reflexivity, agency, and relational markers. These findings indicate changes in observable textual and narrative features and should not be interpreted as direct evidence of corresponding

changes in emotional processing, identity development, or psychological functioning.

The comparison between conditions also highlighted a more complex pattern than simple textual improvement. AI-assisted narratives tended to display greater organization, explicitness, and reflective closure, while the qualitative analysis identified recurrent structures and interpretative trajectories that may indicate a degree of narrative standardization. Greater textual coherence, therefore, cannot be assumed to correspond to greater authenticity. Rather, the findings point to a tension between expressive support and the preservation of individual narrative voice, ambiguity, and authorial control.

From an educational perspective, AI-assisted writing may offer opportunities for making emotions more explicit, reorganizing experience, and comparing alternative ways of expressing the same narrative content. At the same time, these opportunities are accompanied by potential risks, including uncritical acceptance of generated formulations, narrative homogenization, reduced authorial agency, and increasing reliance on algorithmic mediation. The pedagogical value of generative AI may therefore lie less in producing formally improved texts than in creating opportunities for critical comparison between autonomous and AI-assisted writing. Such comparison can support learners in identifying what the system retains, reformulates, adds, simplifies, or displaces.

Within this perspective, DAL may be understood as the capacity to engage with generative systems while maintaining critical awareness of how algorithmic mediation affects emotional language, narrative form, and perceived authorship. Rather than assuming that AI either enhances or diminishes authentic expression, DAL emphasizes reflective ownership of the final text and the ability to evaluate the relationship between personal experience, linguistic representation, and AI-generated reformulation.

Several limitations should be considered when interpreting the findings. First, the purposive sample and the inclusion of two specific participant groups limit the generalizability of the results to other educational, age, and cultural contexts. Second, the study employed a fixed within-participant sequence in which the autonomous narrative always preceded the AI-assisted condition. Consequently, part of the observed differences may reflect order effects, practice, increasing familiarity with the task, or progressive reflection on the narrative topic rather than AI mediation alone. Future studies should therefore adopt randomized or counterbalanced designs.

Third, the study relied on a single generative AI platform, ChatGPT, and the findings cannot be assumed to generalize to other models or interfaces. Moreover, participants constructed different numbers and types of prompts, and prompt quality was not independently assessed. Since the number of prompts was associated with change in text length, future research should examine more systematically how prompting strategies, prompt complexity, and iterative interaction influence narrative outcomes.

Fourth, several narrative indicators, including reflexivity, agency, relational complexity, authenticity, and authorship, necessarily involve interpretative judgments. Although the study sought to increase transparency through explicit operational definitions and reflexive thematic analysis, these constructs should not be treated as fully objective properties of the texts. In particular, authenticity and authorship were explored through participants' metacognitive reflections and researchers' qualitative interpretation rather than through validated quantitative measures.

Fifth, the study did not include longitudinal follow-up. It is therefore not possible to determine whether the observed textual changes persist over time, influence subsequent autonomous writing, or affect participants' longer-term

relationship with AI-mediated writing. Future longitudinal research should investigate whether repeated use of generative AI contributes to the development of narrative and emotional expression or, alternatively, increases dependence on algorithmically generated formulations.

Taken together, the findings support a cautious view of generative AI as a potentially useful educational mediator rather than as a substitute for personal narrative agency. Its value appears to depend on how the interaction is pedagogically framed and on whether learners are encouraged to critically examine, revise, and take ownership of AI-assisted output. Future research should therefore combine stronger experimental designs, multiple AI platforms, systematic analysis of prompting strategies, longitudinal observation, and more specific measures of perceived authenticity and narrative authorship.

References

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

Bruner, J. (1990). *Acts of meaning: Four lectures on mind and culture*. Harvard University Press.

Collaborative for Academic, Social, and Emotional Learning. (2020). *CASEL's SEL framework: What are the core competence areas and where are they promoted?*

Cheng, Y., Fan, Y., Li, X., Chen, G., Gašević, D., & Swiecki, Z. (2025). Asking generative artificial intelligence the right questions improves writing performance. *Computers and Education: Artificial Intelligence, 8*, Article 100374. https://doi.org/10.1016/j.caeai.2025.100374

Chiu, T. K. F., Ahmad, Z., Ismailov, M., & Sanusi, I. T. (2024). What are artificial intelligence literacy and competency? A comprehensive framework to support them. *Computers and Education Open, 6*, Article 100171. https://doi.org/10.1016/j.caeo.2024.100171

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

Draxler, F., Werner, A., Lehmann, F., Hoppe, M., Schmidt, A., Buschek, D., & Welsch, R. (2024). The AI ghostwriter effect: When users do not perceive ownership of AI-generated text but self-declare as authors. *ACM Transactions on Computer-Human Interaction, 31*(2), Article 25, 1–40. https://doi.org/10.1145/3637875

Gross, J. J., & John, O. P. (1997). Revealing feelings: Facets of emotional expressivity in self-reports, peer ratings, and behavior. *Journal of Personality and Social Psychology, 72*(2), 435–448. https://doi.org/10.1037/0022-3514.72.2.435

Hancock, J. T., Naaman, M., & Levy, K. (2020). AI-mediated communication: Definition, research agenda, and ethical considerations. *Journal of Computer-Mediated Communication, 25*(1), 89–100. https://doi.org/10.1093/jcmc/zmz022

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Immordino-Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education, 1*(1), 3–10. https://doi.org/10.1111/j.1751-228X.2007.00004.x

Kring, A. M., Smith, D. A., & Neale, J. M. (1994). Individual differences in dispositional expressiveness: Development and validation of the Emotional Expressivity Scale. *Journal of Personality and Social Psychology, 66*(5), 934–949. https://doi.org/10.1037/0022-3514.66.5.934

Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open, 6*, Article 100174. https://doi.org/10.1016/j.caeo.2024.100174

Li, B., Tan, Y. L., Wang, C., & Lowell, V. (2025). Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence, 9*, Article 100445. https://doi.org/10.1016/j.caeai.2025.100445

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

McAdams, D. P. (2001). The psychology of life stories. *Review of General Psychology, 5*(2), 100–122. https://doi.org/10.1037/1089-2680.5.2.100

Park, J. (2025). A systematic literature review of generative artificial intelligence (GenAI) literacy in schools. *Computers and Education: Artificial Intelligence, 9*, Article 100487. https://doi.org/10.1016/j.caeai.2025.100487

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications.

Riessman, C. K. (2008). *Narrative methods for the human sciences*. SAGE Publications.

Roe, J., & Perkins, M. (2024). *Generative AI and agency in education: A critical scoping review and thematic analysis* [Preprint]. arXiv.

Shi, Y., Yu, K., Dong, Y., & Chen, F. (2026). Large language models in education: A systematic review of empirical applications, benefits, and challenges. *Computers and Education: Artificial Intelligence, 10*, Article 100529. https://doi.org/10.1016/j.caeai.2025.100529

Silvera, D. H., Martinussen, M., & Dahl, T. I. (2001). The Tromsø Social Intelligence Scale, a self-report measure of social intelligence. *Scandinavian Journal of Psychology, 42*(4), 313–319. https://doi.org/10.1111/1467-9450.00242