IJILT
43,4

Mapping AI feedback under emotional prompts: a sentiment-topic-network approach

Li Xiangming and Wei Wei
Faculty of Applied Sciences, Macao Polytechnic University, Macao SAR, China

Received 23 April 2025
Revised 30 November 2025
14 March 2026
Accepted 18 April 2026

Abstract

Purpose – While generative artificial intelligence (AI) is increasingly used to generate feedback on student writing, little is known about how emotional prompts affect the structure of such feedback. This study aims to examine how large language models respond to positive, neutral and negative prompts in postgraduate scientific writing, focusing on how emotional cues shape sentiment polarity, thematic content and structural composition in AI-generated feedback and how these patterns inform learning-relevant feedback architectures.

Design/methodology/approach – An integrated sentiment-topic-network framework of VADER sentiment analysis, Latent Dirichlet allocation topic modelling and epistemic network analysis (ENA) was used to analyse 198 AI-generated feedback messages on 66 postgraduate scientific reports. Sentiment analysis classified emotional polarity, topic modelling identified thematic clusters within which sentiments were expressed, and ENA modelled co-occurrence patterns among cognitive, affective and dialogic codes to characterize feedback architectures under positive, neutral, and negative prompts.

Findings – Across all conditions, positive sentiment dominated even when prompts were neutral or negative, with only negative prompts causing substantial increases in negative codes. Topic–sentiment patterns showed that certain feedback themes were selectively intensified by emotional tone. ENA identified three distinct feedback architectures: supportive-integrative networks under positive prompts, fragmented–minimal networks under neutral prompts, and critical–precision networks under negative prompts. Together, these patterns indicate that emotions modulate not only what feedback is generated but also how cognitive and affective elements are integrated.

Research limitations/implications – Data were drawn from 66 postgraduate students at a single institution, yielding a small, context-specific convenience sample. Findings are therefore interpreted as exploratory patterns rather than statistically generalizable estimates, with analytic generalization aimed at comparable science, technology, engineering and mathematics (STEM)-oriented postgraduate contexts. We did not measure achievement gains or long-term retention. Future research should examine more complex emotional contexts and prompt designs, use mixed-method approaches that include qualitative analysis of emotional nuance and empirically test the three hypotheses about revision quality, perceived usefulness, trust and emotional safety.

Practical implications – Educators can embed prompt templates corresponding to supportive–integrative, fragmented–minimal and critical–precision architectures into course guidelines and teach students when and how to invoke them. Positive prompts can be used to normalize difficulty and sustain engagement in early drafting, with brief episodes of critical–precision feedback introduced once drafts are coherent. Developers of institutional AI tools can implement visible "feedback mode" selectors that map onto these architectures, allowing teachers to align system behaviour with course-level pedagogy rather than relying on a single default.

Originality/value – This study advances an integrated Sentiment–Topic–Network framework that links sentiment polarity, thematic focus and network-level co-occurrence patterns in AI-generated feedback. It empirically characterizes a robust positivity bias in large language model feedback under positive, neutral, and negative prompts and identifies three feedback architectures – supportive-integrative, fragmented–minimal and critical–precision – relevant for emotionally tuned feedback design. By translating these architectures into an adaptive model and testable hypotheses, the study offers a reusable lens for analysing emotionally primed AI feedback in postgraduate STEM-oriented learning contexts.

Keywords AI feedback, Emotional prompts, Epistemic network analysis, Sentiment analysis, Topic modelling

Paper type Research article

Funding: This work was supported by Macao Polytechnic University Research Grant (award no: RP/FCA-07/2025).

Competing interests statement: The authors declare no competing or conflicting interests.

The International Journal of Information and Learning Technology
Vol. 43 No. 4, 2026
pp. 388-415
Emerald Publishing Limited
e-ISSN: 2056-4899
p-ISSN: 2056-4880
DOI 10.1108/IJILT-04-2025-0121

1. Introduction

Generative AI such as GPT-4.0 and GPT-4o has revolutionized artificial intelligence's capacity to generate human-like text. These advancements have had a significant impact on a variety of domains, particularly sentiment analysis, where machines can now replicate the subtleties of human emotion (Kit and Mokji, 2022). Emotion-driven prompts for text generation are increasingly incorporated into AI-driven communications to enhance interactivity and personalization (Kumawat *et al*., 2021). In particular, emotional cues have emerged as an important component of sentiment analysis. Through the use of emotion-induced prompts, generated content has become more relevant and more applicable across a variety of emotional and psychological contexts (Vickneswaran *et al*., 2020). Recent work also examines the behaviour of AI chatbots in response to emotional prompts and has shown that advanced models like ChatGPT-4 can exhibit response patterns that are influenced by emotional cues (Zhao *et al*., 2024).

Prior research has largely focused on the technical aspects of language model training and their application in neutral settings without addressing the emotional depth that can be induced by tailored prompts. Recent studies have demonstrated superior performance in text generation with controlled sentiment influenced by structured data sets and predefined emotional cues using multimodal features and generative models (Mao *et al*., 2023; Zhu *et al*., 2022). Further advancements have focused on the role of data augmentation and neural network architectures in improving the sentiment analysis capabilities of these models (Aini *et al*., 2023).

Despite these advances, little is known about how generative AI dynamically incorporates emotion-laden cues when producing feedback in authentic educational environments. This creates a critical conceptual gap: it remains unclear how emotion-eliciting prompts interact with textual feedback from generative AI, especially when differences in emotional valence and intensity are dynamically introduced. This interaction is crucial for determining whether AI feedback architectures are affected systematically by emotional framing and how such variations influence learning (Zhang *et al*., 2023). Methodologically, previous studies have seldom combined sentiment analysis, topic modelling, and relational network analysis within a single framework in order to capture multi-layered effects in emotion-triggered AI responses. This gap limits our understanding of how emotional tone influences both the linguistic composition and the structural organization of AI-generated feedback. Furthermore, few studies have addressed these dynamics at a contextual level, where the relational structure of feedback-the way comments relate across affective and cognitive dimensions-is crucial to learner engagement. These intertwined gaps motivate the present study.

Accordingly, this study examines generative-AI feedback produced under emotion-laden prompts in order to fill these conceptual, methodological, and contextual gaps. We conceptualize emotional prompting as a design lever capable of reconfiguring how cognitive, dialogic, and affective feedback elements are integrated rather than treating it as a purely technical control for feelings. This perspective is operationalized through the Sentiment-Topic-Network (STN) framework, which combines sentiment analysis, topic modelling, and Epistemic Network Analysis (ENA) to reveal interconnected patterns of emotional and cognitive feedback. This framework allows for the analysis of not only surface sentiment, but also internal co-occurrence networks of key feedback features as a result of emotional prompts. Building on this framework, this study aims to advance pedagogical understanding of emotionally induced AI feedback as opposed to developing new algorithms.

The study makes three interrelated contributions. Firstly, it proposes a Sentiment-Topic-Network (STN) framework for analysing sentiment polarity, thematic focus, and network-level co-occurrence in AI-generated feedback. Second, it provides empirical evidence of a systematic positivity bias and demonstrates how this bias can be reconfigured under different emotional framings. Third, it translates these patterns into design-relevant insights for learning by demonstrating how emotional prompts restructure the coupling of cognitive clarity, dialogic responsiveness, and affective tone-factors that influence learners' motivation,

perceptions of the usefulness of feedback, and their willingness to act. Together, these contributions position emotional prompt design as a critical variable in AI-mediated learning and provide a framework for examining emotionally adaptive AI feedback in broader contexts.

2. Literature review

2.1 Types of generative AI-assisted feedback in response to prompts

Generative AI feedback refers to the outputs produced in response to inputs, specifically prompts. Feedback generated by AI can be broadly categorized according to accuracy, relevance, emotional resonance, and coherence (Korinek, 2023; Lee and Park, 2023). These feedback categories contribute to understanding how effectively generative AI can interpret and respond to prompts. These dimensions are interpreted not only as technical properties of system output in educational settings, but also as pedagogically meaningful features that influence how learners perceive, interpret, and implement feedback within broader traditions of formative, dialogic, and affect-sensitive feedback. For example, the study of AI-assisted generative feedback revealed that the model was capable of interpreting prompts and producing responses that are consistent with the desired attributes or behaviours (Ruotsalo *et al*., 2023).

A fundamental aspect of generative AI-assisted feedback is accuracy and relevance. AI models tend to produce precise and relevant responses when given clear and unambiguous instructions. Kasneci *et al*. (2023) demonstrated that explicit and precise prompts improve generative AI output accuracy and relevance, particularly in the context of education. Therefore, prompt design is crucial to eliciting the desired responses from artificial intelligence systems.

Another critical aspect is emotional resonance, especially in applications such as customer service and therapy in which the emotional tone of the response has a profound impact on client satisfaction and effectiveness. As reported by Rashkin *et al*. (2019), prompts labelled with specific emotional cues elicit responses that are more closely aligned with the desired emotional state. As a result, eliciting a more robust response enhances user engagement. Furthermore, Ruotsalo *et al*. (2023) discovered that by incorporating feedback mechanisms that allow AI models to interpret emotional cues, text generation can be even more powerful in achieving emotional resonance.

The third aspect of AI-assisted feedback is coherence, which refers to how well AI-generated responses follow a logical flow and maintain clarity. Both aspects of logic flow and clarity are crucial for maintaining user engagement and encouraging natural conversation between humans and AI. Research has shown that training dialogue models on large and diverse conversational datasets leads to improved topic consistency and perceived quality of multi-turn interactions (See *et al*., 2019). A further study examined the role of memory effects in enhancing neural coherence (Xu and Jin, 2020). Their findings provided insight into optimizing feedback mechanisms to enhance neural network coherence.

2.2 Impact of emotional prompt design on generative AI-assisted feedback

Recent research has found that emotional prompt design has significantly influenced generative AI-assisted feedback. Effective prompt design can enhance the emotional sensitivity and accuracy of content generated by generative AI models. This improvement makes these models more suitable for applications that require a comprehensive understanding of emotions.

For instance, incorporating emotional information into prompts has been found to enhance the truthfulness and usefulness of AI-generated responses. Accordingly, psychological principles may be strategically applied in prompt design to increase emotional resonance and accuracy in intelligent generative systems (Zhou *et al*., 2018). Designing prompts thoughtfully can also assist in aligning generative AI with user intentions. When models are fine-tuned with

human feedback including prompts, toxic outputs are reduced and overall accuracy is improved. This illustrates the importance of emotional prompts and user feedback in optimizing AI performance (Ouyang *et al*., 2022). In educational settings, emotional prompts also play a significant role in enhancing learning and motivation, which provided further evidence of the need for intentional prompt construction (Li *et al*., 2023). From a learning perspective, this issue is closely related to feedback literacy, social-emotional support and learner-feedback interaction, since the value of AI-generated feedback depends not only on the quality of output, but also on learners' capacity to interpret, trust, and utilize it.

Additionally, prompt design is able to mitigate biases within generative AI, especially

When it comes to sentiment analysis and emotion recognition. Selecting prompt templates and label words carefully is crucial for reducing bias in AI models (Mao *et al*., 2023). Further, the use of a "Devil's Advocate" approach in prompt design can enhance the robustness and balance of AI-assisted group decision making (Chiang *et al*., 2024). As a result, this approach emphasizes the importance of strategic prompt development in order to prevent unintended consequences and to facilitate effective decision-making.

2.3 Evaluating generative AI-assisted content

Evaluation of generative AI-assisted text requires an assessment of several attributes, such as coherence, relevance, and factual accuracy. Previous studies have emphasized the need for reliable and contextually appropriate AI outputs. For instance, Ji *et al*. (2023) emphasized the importance of incorporating fact-checking mechanisms in order to enhance the reliability of generative AI-based content. Researchers have also demonstrated that it is effective to differentiate content created by computer-generated systems and human-generated systems (Nguyen *et al*., 2023).

Topic and sentiment analysis have emerged as critical issues for improving the evaluation of AI-generated texts. A comprehensive understanding of the sentiments and topics conveyed by AI models is essential for applications requiring context awareness and emotional resonance. Various methods for sentiment analysis have been proposed, including machine learning and statistical language models, with the purpose of detecting sentiments in user-generated content (Höpken *et al*., 2017). As illustrated in Sun *et al*. (2023), generative AI systems might perform iterative reasoning and correct each other to enhance sentiment analysis. The system was based on having multiple generative AI systems reasoning and correcting each other iteratively.

In order to ensure that AI models produce high-quality, emotionally resonant, and contextually appropriate content, researchers focus on topic and sentiment analysis along with coherence, relevance, and factual accuracy.

Despite substantial advancements in the field, a significant gap remains in the integrated understanding of how different emotional prompt designs affect generative AI-assisted feedback (e.g. Ficler and Goldberg, 2017; Hu *et al*., 2017; Keskar *et al*., 2019). Conceptually, it remains unclear how specific emotional cues embedded within prompts influence sentiment, thematic emphasis, and dialogic qualities of generated text. Methodologically, most previous studies have examined sentiment and topic layers in isolation without connecting these layers to the internal co-occurrence structure of feedback features. Contextually, while neutral prompts in general-purpose settings have been shown to generate texts with a positive sentiment bias (Yang *et al*., 2020), empirical studies situated in educational feedback settings remain scarce.

At a more detailed level, many existing studies focus on manipulating sentiment in non-educational corpora, evaluating AI outputs primarily on their accuracy or user satisfaction, or analysing human–human or human-AI discourse structures without explicitly modelling the emotional cues embedded in prompts (Ficler and Goldberg, 2017; Gilardi *et al*., 2023). As a result, we know relatively little about how emotional prompts simultaneously shape (1) the polarity and intensity of AI feedback, (2) the thematic emphases associated with different

emotional stances, and (3) the internal coupling of cognitive, affective, and dialogic dimensions in feedback on student work (Shute, 2008).

Furthermore, the structure of feedback has rarely been considered in models of AI-human interaction in education. A majority of ENA studies have mapped epistemic moves in human teams or human-AI collaborative activities (Ba *et al*., 2024; Fernandez-Nieto *et al*., 2021; Viberg *et al*., 2024), but no study has attempted to characterize the network structure of AI-generated feedback itself under systematically varied emotional stimuli. A number of learning analytics and discourse analysis studies have examined interaction structure and instructional dialogue, but few have integrated emotional prompt variation, thematic modelling, and structural coupling into one framework for analysing AI-generated feedback (Banihashem *et al*., 2022; Smit *et al*., 2023). ENA combined with sentiment-topic analysis allows us to conceptualize AI feedback as a networked architecture embodying different feedback philosophies (e.g. supportive versus corrective, dialogic versus fragmented) under different emotional frames, which motivates the integrated framework we adopt in this study.

To address these conceptual, methodological, and contextual gaps, this study applies a Sentiment-Topic-Network (STN) framework to generative AI feedback elicited by positive, neutral, and negative prompts. Three analytically linked layers including sentiment analysis, topic modelling, and Epistemic Network Analysis (ENA) together model emotional tone, thematic focus, and structural organization in AI feedback. This threefold structure aligns the STN framework conceptually with learning science concerns, capturing affective stance, substantive guidance, and pedagogical feedback organization. The approach goes beyond using existing tools solely as analytic instruments to generate new empirical insights into how prompt tone reconfigures feedback architecture in educational settings.

Therefore, the following research questions were proposed:

(1) How do positive, neutral, and negative prompts differentially shape (a) sentiment polarity, (b) topic configuration, and (c) structural co-occurrence patterns among cognitive, affective, and dialogic feedback attributes in generative AI-assisted feedback?

(2) To what extent do the emotions represented in generative AI-assisted feedback align with the intended emotional triggers embedded in the corresponding prompts?

3. Methodology

3.1 Research philosophy

This study adopts a pragmatic and exploratory research philosophy. Instead of testing a predefined causal model of learning outcomes, we explore how emotional prompts shape the properties of AI-generated feedback and present the tool as a designable tool within an authentic postgraduate teaching context. Our objective is to generate practical, theory-informed insights that will guide subsequent design and experimental work on educational technology that is emotionally aware.

Within this perspective, we examined how emotional cues embedded in prompts have an effect on both the tone and content of AI-generated feedback, as well as students' emotional reactions to these variations. The study aimed to demystify how generative feedback is produced and how such feedback should be interpreted as opposed to encouraging an excessive reliance on artificial intelligence.

A pragmatic approach also included careful consideration of ethical issues related to the use of generative AI tools in the classroom, such as ChatGPT. Several ethical concerns have also been raised regarding the integration of generative AI tools in educational settings, such as ChatGPT. Students' rights to control their own academic content may be violated when instructors input student work into artificial intelligence systems without obtaining clear consent (Williams, 2024). A second concern is the possibility of compromising the privacy and

intellectual property rights of students who upload their works to commercial, cloud-based platforms, especially when data may be stored or reused without transparency (Abdullah *et al*., 2024; Ahmed, 2024). As a third consideration, relying on AI to generate feedback may undermine the authenticity of teacher-student interaction and lessen the instructor's role in supporting student learning (Cowling *et al*., 2023). Furthermore, repeated exposure to AI-generated responses may lead students to rely on these tools without critically evaluating their quality, particularly when such feedback lacks human insight or subject-matter understanding (Nikolopoulou, 2024; Ocampo *et al*., 2023).

This study was designed with the intent of avoiding each of these ethical pitfalls. Firstly, all participants provided informed consent and participated in the study as part of their regular classroom activities. Students' work was not used or submitted without their knowledge or permission. Secondly, students manually entered their own writing ranging from personal coursework to published abstracts and publicly presented lab reports—into the AI platforms. Thirdly, the use of AI was explicitly described as exploratory and non-evaluative: students applied emotion-laden prompts (positive, neutral, negative) to generate various forms of feedback. Fourthly, the activity was designed to promote critical thinking about the nature, limits, and implications of AI-generated feedback, not to replace instructor interaction.

3.2 Research approach

In terms of the research approach, the study is exploratory and predominantly inductive in nature. Our study begins with empirical patterns found in a corpus of AI-generated feedback, which are used to refine a Sentiment-Topic-Network (STN) framework and to derive hypotheses about the "feedback architectures" of emotionally tuned artificial intelligence systems. An alternative approach to evaluating structural models in higher education focuses on analysing the sentiment distributions, topic configurations, and network structures as emergent properties that inform the development of theories about emotional prompt design and AI-mediated feedback.

We conducted a corpus-based analysis of AI-generated feedback in an authentic postgraduate learning environment. 66 postgraduate students from two intact classes participated in an experimental activity conducted over 2 weeks at a single research university in the People's Republic of China. While this study's broader pedagogical context involved developing students' scientific writing, its specific focus was investigating how the large language model (LLM) responded to emotionally framed prompts during feedback.

A key aspect of this paper is that the unit of analysis is the AI-generated feedback text rather than the students themselves. Students' scientific reports provided realistic, discipline-specific input that ensured ecological validity, but the outcome variables (sentiment, topics and structural co-occurrence patterns) were properties of AI feedback. Thus, rather than measuring learning gains or performance changes directly, this study contributes to the learning field by characterizing the emotional and structural features of AI feedback.

3.3 Research method: integrated sentiment-topic-network framework

We employ a multi-layer, corpus-based method of text analysis. An integrated Sentiment-Topic-Network framework was used to analyse the dataset, which captures three complementary dimensions of AI-generated feedback: (1) affective polarity, (2) thematic focus, and (3) structural organization. By combining sentiment analysis, topic modelling, and Epistemic Network Analysis (ENA) sequentially, we were able to construct a multilayered representation of how emotions shape AI feedback. Sentiment analysis and topic modelling indicate emotional valence and thematic emphasis, whereas the network layer examines how evaluative, cognitive, and affective factors interact within feedback and provides insights not available from sentiment-topic analysis alone. The framework, however, also involves analytic trade-offs since it prioritizes interpretable cross-layer comparison over fine-grained emotional modelling, platform-specific analysis, or causal explanation, and these trade-offs

define its boundary conditions. This multi-layered design is intended to capture both surface-level and deeper structural characteristics of AI feedback that are theoretically relevant to learners' evaluation and support experiences.

Specifically, sentiment analysis provides a first-layer quantification of emotional valence in text. The resulting sentiment classes are then utilized to define and validate the affective codes for the ENA. The topic modelling technique provides a second layer for identifying the thematic clusters within which sentiments are expressed and for refining the cognitive dimensions (e.g. clarity, coherence, accuracy) used in ENA. Finally, ENA adds a structural layer that models how these sentiment- and topic-informed feedback dimensions co-occur within and across emotional conditions, extending previous sentiment-topic research of affective-conceptual interactions with network-based modelling of feedback structure (Fatima-Zahrae and Wafae, 2021; Ohmura *et al*., 2014; Osmani *et al*., 2020).

3.3.1 Sentiment analysis (first layer). To begin with, all 198 AI-generated feedback texts were pre-processed using the Natural Language Toolkit (NLTK). A series of steps were taken to standardize and de-noise the data, including lowering the case, tokenizing, removing stopwords, and generating word stems. Each full feedback text was then processed using the Valence Aware Dictionary and Sentiment Reasoner (VADER). The VADER algorithm produces a compound sentiment score ranging from −1.0 (strongly negative) to +1.0 (strongly positive) based on the combination of lexicon-based and rule-based features. VADER scores were calculated in this study in order to determine whether each AI-generated feedback text was positive (a compound score above 0.05), neutral (between −0.05 and 0.05), or negative (a compound score below −0.05).

This automated process resulted in an efficient and consistent method for the classification of sentiments at a large scale. Considering that VADER is an established and well-validated text analysis tool with proven effectiveness across a wide range of text domains (Telmo *et al*., 2024), the exploratory nature of the study warranted the application of this proven and mature methodology (Taj and G.S., 2024). While lexicon-based sentiment scoring may not fully capture subtle emotional nuances in educational and academic feedback discourse, it remains methodologically appropriate for the present analysis. We focus on broad polarity patterns across prompt conditions which support valid comparison. These three sentiment classes served as a basis for constructing and calibrating the three affective codes in the ENA coding scheme (Affective_Positive, Affective_Neutral, Affective_Negative). This provided an empirically grounded mapping between the automated polarity scores and the human-coded affective categories later used for network analysis.

3.3.2 Topic modelling (second layer). The second phase focused on analysing the thematic structure of AI-generated feedback using Latent Dirichlet Allocation (LDA). In conjunction with the pre-processing described above, a bag-of-words document–term matrix was created using the Gensim library, which enabled the LDA training procedure to be based upon the essential document–term matrix. By comparing models with varying number of topics, we found an interpretable topic solution using coherence and perplexity scores, as recommended by Blei *et al*. (2003) and Röder *et al*. (2015). A final model identified five topic clusters for negatively prompted texts and four clusters for neutrally and positively prompted texts.

We identified topics by manually reviewing excerpts of representative feedback and analysing their top-weighted keywords. The interpretation was based on coherence, lexical consistency, and alignment with feedback attributes (e.g. clarity, structure, tone). These outputs were not used simply as descriptive tools, but directly informed operationalization and refinement of the cognitive dimensions in the ENA coding framework. The topics that emphasize "structure", "logical flow", or "sections" were evidence of the relevance of a Coherence dimension, whereas topics highlighting "clarity", "language", or "expression" supported the Clarity dimension. Integrating topic modelling and sentiment analysis enabled a multifaceted analysis of how emotional triggers shaped AI-generated feedback by situating polarity within specific thematic foci.

This combined sentiment–topic approach was selected for this study based on its maturity and reliability in corpus-based research (Taj and G.S., 2024).

3.3.3 Epistemic Network Analysis (third layer). In the third phase, Epistemic Network Analysis (ENA) was employed to examine the conceptual structure of AI-generated feedback beyond sentiment polarity and topic distribution. While sentiment analysis captures the emotional valence of feedback, and LDA identifies dominant theme clusters, ENA adds a relational dimension by revealing how different feedback features occur within and across local segments. The structural perspective is particularly useful for analysing the logic and cohesion of feedback, which failed to be adequately captured by frequency- or polarity-based approaches alone. By visualising the co-occurrence patterns of key evaluative elements, ENA facilitates a deeper understanding of the tone and pedagogical design of AI-generated instructional responses (Fernandez-Nieto *et al*., 2021; Viberg *et al*., 2024). Within the integrated Sentiment-Topic-Network framework, the ENA models therefore draw on affective and cognitive codes that are empirically calibrated by the preceding sentiment analysis and topic analysis, rather than defined purely *a priori*.

We developed a coding framework consisting of eight binary-coded dimensions, informed by literature on effective feedback and natural language generation. Four cognitive attributes are coded: Accuracy (correctness and specificity), Relevance (alignment with the student input), Clarity (linguistic simplicity), Coherence (logical organization), along with dialogic attribute about explicit connection to prompt elements (Prompt-Responsiveness). Three affective tone variables reflected the emotional context of the language: Affective_Positive, Affective_Neutral, and Affective_Negative. The affective codes were calibrated using VADER-derived sentiment classes, and the cognitive codes were refined based on the LDA topics (e.g. structural organization, linguistic clarity, task alignment). This coding scheme ensured both theoretical coverage and methodological alignment across the analytical layers (Adie *et al*., 2018; Naz and Robertson, 2024).

All coding was completed independently by two trained human raters with expertise in AI-assisted educational feedback and discourse analysis. Each feedback text was segmented into stanzas, and codes were applied as binary variables at the stanza level, with 1 indicating the presence of the corresponding feature and 0 indicating its absence. Inter-rater reliability assessed by Cohen's kappa was 0.87, which indicates significant coding consistency between the two raters, The final coding scheme was finalized through a process of discussion and consensus.

Affective tone was determined by assigning each feedback stanza one of three binary codes. For each feedback stanza, exactly one of the three codes—Affective_Positive, Affective_Negative, or Affective_Neutral—was set to 1 based on its dominant sentiment orientation, while the remaining two were set to 0. The Affective_Positive = 1 designation indicates praising, encouraging, or affirming language; the Affective_Negative = 1 designation indicates critical evaluations; and the Affective_Neutral designation reflects neutral or objective language. The codes were applied independently based on the characteristics of the tone. This structure maintained emotional distinctions in ENA while enabling condition-specific comparisons of network architectures (Herb and Lloyd, 2024; Kuzminykh *et al*., 2024).

We used the ENA Web Tool to generate separate networks for positive-, neutral-, and negative-prompt conditions. The unit of analysis was each participant's AI-generated feedback within a given emotional condition, and conversation boundaries were defined by condition (e.g. all negative-prompt feedback for a given participant). A moving window size of one stanza was selected to capture tight co-occurrences between adjacent codes. In the resulting ENA network, nodes represent feedback dimensions, while edges represent the strength of co-occurrence among nodes. Thicker edges indicate stronger co-occurrence between nodes.

3.3.4 Analytic integration across the three layers. We interpreted the results across the three analytical layers rather than in isolation in order to maintain methodological coherence.

For instance, the observed positivity bias in VADER scores provided the baseline against which we interpreted the more evaluative, accuracy-centred co-occurrence networks that emerged under negative prompts. Similarly, topics emphasizing structural aspects such as organization and method description helped explain why coherence and accuracy formed stable hubs in the ENA networks across all condition. Within this integrated Sentiment-Topic-Network framework, we have been able to treat the ENA maps not as stand-alone visualizations, but as structural corroborations or further refinements of patterns initially suggested by the sentiment and topic analyses.

3.4 Research strategy

Research strategy involves a quasi-experimental corpus-based design incorporated into a regular postgraduate writing course. An exploratory classroom activity was conducted over a period of two weeks with two intact classes at a research university located in the P.R.China. The broad pedagogical context of this study involved developing students' scientific writing, while the specific purpose of this study was to investigate whether large language models (LLMs) responded to emotionally framed prompts when providing feedback to students.

Initially, the participants prepared original scientific reports prior to interacting with the generative AI. These reports provided input to the LLMs and served as a baseline for the generation of subsequent feedback based on different emotional triggers.

Subsequently, the prompts were constructed to simulate supervisors with differing communication styles, such as harsh, neutral, and friendly, which reflect negative, neutral, and positive emotional tones, respectively. We kept the informational focus constant throughout all three versions by addressing clarity, structure, and methodological description while systematically altering the wording and evaluative stance to convey different emotional connotations.

The three prompts were submitted separately for each student report to the AI platform, and corresponding feedback texts were collected. AI-generated feedback instances were tagged with the ID of the source report and the emotional condition of the prompt (positive, neutral or negative). This tagging produced a structured corpus of feedback for subsequent analysis. The above procedure ensured that each original text was associated with one AI response for each emotion.

This design simulates a scenario increasingly encountered in instructional practice, in which educators use language models to provide formative feedback to students. Recent surveys and practitioner reports indicate that instructors are experimenting with AI tools in support of feedback processes, particularly in environments with high enrolments or time constraints (Cowling *et al*., 2023; Pang *et al*., 2024). We developed tones of prompts that mimic the three typical supervisory styles of supportive (positive), neutral, and critical (negative). Thus, our design addresses a question directly relevant to the learning field about how does emotional framing affect the types of evaluative messages students receive from AI-based feedback.

3.5 Data collection instruments and source

The data collection process drew on authentic student texts, emotion-laden prompts, and multiple LLM-based writing assistants. A total of 66 independent scientific reports were used in the experiment, each limited to 150 words and written in accordance with academic writing standards. The texts were drawn from students' existing work, such as course assignments, journal abstracts, and small summaries presented in lab groups. The texts were carefully selected to maintain objectivity and neutrality so that the results of the study would not be affected by pre-existing emotional content. Thus, any observed emotional variation could be attributed primarily to the emotional prompts applied to the AI, rather than by the affective content in the source texts.

For each report, we created three prompts that differed only in their emotional tone (positive/friendly, neutral, and negative/critical) while maintaining comparable informational content. The prompts were framed as messages from a hypothetical supervisor commenting on the same piece of student writing ranging from supportive and constructive to blunt and critical. Table 1 summarizes the core prompt templates and their emotional classifications. Although the three conditions differed in evaluative language and affective stance, all three highlighted similar aspects of student work, including clarity, structure and methodological description.

These neutral academic texts were used as consistent inputs to several generative AI platforms (including ChatGPT, Copilot, ChatGLM, Kimi, and Wenxinyiyan). The present analysis treats these platforms as instances of advanced LLM-based writing assistants, aggregating their outputs rather than considering platform-by-platform comparisons. By including mixed-platform usage in contemporary academic settings, the study will have a wider ecological relevance, and model heterogeneity will be treated as a boundary condition rather than a base for platform-specific comparisons. A total of 198 AI-generated feedback texts were generated by applying three emotion-laden prompts to every report (66 reports x 3 prompts). Examples of AI-generated responses include:

(1) "*Excellent clarity in your methods section. Your logical flow really stands out—great job!"* (Positive prompt)

(2) "*The methods are described and appear to align with the study's aim."* (Neutral prompt)

(3) "*The explanation of your methodology lacks precision and requires clearer structure."* (Negative prompt)

Standardized, emotionally neutral input texts allowed us to isolate the influence of emotional prompting on generative AI output while maintaining high ecological validity. The resulting corpus of 198 feedback texts provides a controlled yet realistic basis for examining how emotional prompts reshape the affective, thematic, and structural properties of AI-assisted feedback.

3.6 Population

The population of interest consists of postgraduate students enrolled in scientific and engineering programmes at a research-intensive university in P.R. China who are developing their academic writing skills in English.

A total of 66 postgraduate students (M = 24.23 years, SD = 2.15) participated in this study. The sample consisted of 41 males and 25 females from a range of scientific and engineering disciplines. They participated in an English-medium writing course focused on research writing in which they were routinely required to write abstracts, short reports,

Table 1. Prompts indicating different emotional stimuli

Original version of research progress report for weekly lab meeting *Questions to raise in the chatbot*	Copy paste the original report *Answers generated by chatbot*
1.1. You are my supervisor and you are going to evaluate my lab reports on the following criteria: *indicating Neutral*	*1.1* *Copy Paste from Chatbot*
1.2. You are my supervisor who is very harsh and critical. And you are going to evaluate my lab reports on the following criteria: *indicating Negative*	*1.2* *Copy Paste from Chatbot*
1.3. You are my supervisor who is very considerate and supportive. And you are going to evaluate my lab reports on the following criteria: *indicating Positive*	*1.3* *Copy Paste from Chatbot*

and laboratory summaries, providing an authentic context for examining AI-assisted feedback on disciplinary writing.

3.7 Sampling technique and sample

An intact-class, non-probabilistic sampling strategy was adopted. Two intact classes from the postgraduate writing course were invited to participate, and all students who gave informed consent were included in the study. The sample ($N = 66$) can therefore be characterized as a convenience sample drawn from intact course cohorts, rather than a random sample drawn from a wider postgraduate population.

This sampling technique aligns with the exploratory and context-specific aims of the study. The objective of the study was to obtain a rich, ecologically valid corpus of AI feedback based on realistic postgraduate writing tasks, rather than to estimate demographic parameters with statistical generalizability. We therefore strive for analytic generalization to comparable postgraduate, STEM-oriented contexts rather than broad population-level conclusions.

A Sentiment-Topic-Network framework is used in this study as an exploratory, theory-building approach, as it is based on 66 postgraduate students from a single research university along with 198 AI-generated feedback texts drawn from their writing. This study does not claim statistical generalizability for all learners, institutions, or generative AI systems, but rather seeks analytic generalization to similar STEM-oriented postgraduate education contexts in which instructors and students generate formative feedback through large language models.

4. Results

4.1 Sentiment distribution across emotionally-induced AI feedback

As illustrated in Figure 1, the study conducted a detailed sentiment analysis on texts generated by a large language model that was prompted with emotional cues varying from positive to negative. Analysis of the sentiment distribution across these texts revealed a distinctive pattern in how the emotional tone of the input prompts influenced the emotional content of the output.

Positive-induced prompts led to a significant majority of 66.4% of sentiments being classified as positive, demonstrating the model's tight alignment with input cues. The remaining consisted of 24.0% neutral sentiments, and only a small percentage of 9.6% in negative sentiments. The results indicate that positive prompts not only dominated sentiment output, but also effectively suppressed negative sentiment.

Alternatively, in neutral-induced texts, the sentiment distribution was dominated by positive sentiments, accounting for 64.6% of the emotional tone. Even without explicit

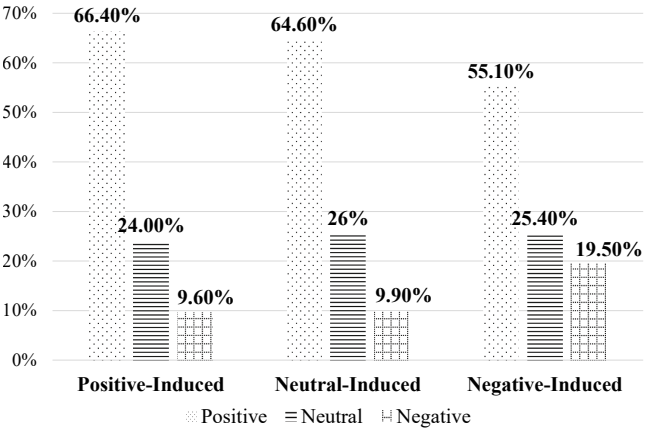

Figure 1. Sentiment analysis results under different emotional prompts. Authors' own work

positive cues, the model produced predominantly positive sentiment outputs, highlighting an inherent positive bias in its response to neutral prompts.

While negative-induced texts displayed a more balanced sentiment distribution, positive sentiments still dominated at 55.1%. Negative sentiments, however, increased significantly to 19.5%, which was twice as high as in the other two types of positive- and neutral-induced texts. The neutral sentiments were observed at 25.4%, which is similar to the sentiments in neutral-induced texts. The result illustrated the model's capability to adapt to the negative tone of the prompts while also maintaining a substantial level of positive output.

In all three types of texts, there is a direct correlation between sentiment distribution and the emotional tone of the prompts used to generate the texts. AI feedback produced a significant share of positive sentiments in both positive- and neutral-prompted texts (66.4 and 64.6%, respectively), which illustrated the tendency to produce positively skewed outputs even when neutral prompts were utilized.

In contrast, texts prompted with negative tone generated an average of 19.5% negative sentiments, almost twice as high as neutral- and positive-prompted texts (9.6 and 9.9%, respectively). This significant increase in negative responses to texts generated as a result of negative prompts suggests that negative prompts have the ability to elicit negative emotional content in texts. The results revealed that generative AI was subject to sentiment shifts induced by prompts.

Across all three conditions, therefore, sentiment distribution only partially aligns with the intended emotional tone of the prompts. Positive prompts produce the best match between prompt valence and output sentiment. By contrast, both neutral and negative prompts elicit predominantly positive feedback, with only modest increases in neutral or negative content. This pattern directly addressed both research questions: emotional prompting affects the polarity of AI-generated feedback, but a persistent positivity bias limits the faithfulness with which neutral and negative prompts are reflected in the resulting texts.

4.2 Topic and sentiment dynamics across emotionally-induced AI feedback

The Latent Dirichlet Allocation (LDA) model was used to extract various topics within generative AI-assisted texts. Each theme was defined by specific keywords that were representative of the core subjects of the feedback. The topics covered in this series generally include linguistic structure, sentiment, and technical quality of responses, reflecting the varied focus of Generative AI when interpreting different prompts.

Based on Table 2, each topic represents a different aspect of the feedback provided by the generative AI. The topics demonstrate how the models handle a range of language generation aspects, from technical accuracy to communicative effectiveness. These insights are essential for understanding the capabilities and limitations of current generative AI-assisted technologies when it comes to producing the contextually appropriate responses.

Further, the distribution of sentiment and topics across AI-generated texts revealed distinct responses, when prompted with emotional cues categorized as positive, neutral and negative. An analysis of sentiment-topic graphs was conducted to determine the intensity and scope of sentiments within specific topics under each type of induction.

The sentiment-topic graph illustrates how emotional prompts influence thematic and emotional content by showing the relationship between topics discussed and sentiments expressed. Previous research has suggested that sentiment and topic modelling can be integrated to improve understanding and visualization of these relationships. An example is Lin *et al*. (2015)'s joint sentiment-topic model that detects sentiment in conjunction with specific topics without labelled data.

Figure 2 illustrated sentiment distributions across three separate graphs, each representing a different emotional prompt (negative, neutral, positive) and covering distinct topics within each graph. Each bar in these graphs shows the frequency of sentiment responses

Table 2. Topic distribution from LDA analysis

Emotion type	Topic item	Content
Negative-Induced	1	Linguistic Quality Assessment: the strengths and weaknesses in linguistic structures
	2	Report Clarity and Structure: clarity and organizational structure of reports
	3	Technical Writing and Terminology: specific language and terms
	4	Critical Evaluation and Analysis: evaluating and critiquing reports or papers
	5	Writing Mechanics and Research Presentation: foundational elements
Neutral-Induced	1	Technical and Scientific Writing Process: language and structural aspects
	2	Clarity and Coherence in Report Writing: well-structured and logically flowing
	3	Effective Communication in Research: effectively communicating research findings
	4	Linguistic Evaluation and Technical Analysis: critical analysis of linguistic strengths and weaknesses
Positive-Induced	1	Linguistic Support and Evaluation: linguistic assessments and the supportive feedback
	2	Research Documentation and Communication: structuring and presentation
	3	Technical Strength and Linguistic Precision: technical robustness and linguistic precision
	4	Structural Clarity in Writing: clarity and structural integrity of written reports or documents

Note(s): Emotional Thematic Generative AI Responses

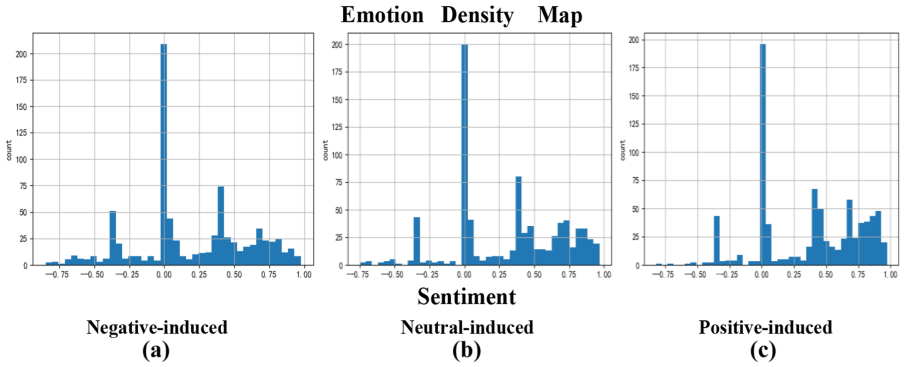

Figure 2. Sentiment-topic distribution under different emotional prompts. Authors' own work

(positive, neutral, or negative) that AI models generated for specific topics under various emotional prompts.

Specifically, graph a (Negative-Induced Prompts) shows an intriguing divergence between expected and actual responses. Topic 1, for instance, despite being prompted with a negative cue, exhibits a significant prevalence of positive responses (95 instances) over negative (70 instances) and neutral (40 instances). This unexpected positivity indicates either a bias in the model towards favourable interpretations or perhaps a positive tone in the topic content itself. Both topics 2 and 3 display similar trends, although topic 3 exhibits a strong lean toward positivity and minimal alignment with the negative prompt. On the other hand, topic 4 exhibits a moderate negative response, but is heavily skewed towards positive sentiment. This reinforces the notion that there is an underlying optimism in the model.

A strong inclination toward positive sentiment is evident in graph b (Neutral-Induced Prompts) even in the absence of an explicit emotional directive, particularly in topics 1, 2, and 3. In contrast, topic 4 demonstrates a more balanced sentiment distribution, displaying substantial examples of each type of sentiment. The pattern in graph b indicates that, while certain topics including topics 1, 2, and 3 showed a natural tendency to skew toward positive sentiments, while topic 4 exhibited a more balanced distribution between positive, neutral, and negative responses. This variation illustrates how specific topic characteristics are independent of the tone of the prompt in influencing sentiment outcomes. Alternatively, under neutral prompts, some topics appear predisposed to evoke positive interpretations, whereas others remain neutral or mixed.

A consistent pattern of positive responses in graph c (Positive-Induced Prompts) across all topics closely aligned with the intended emotional prompt. In particular, topic 3 exhibits a highly skewed sentiment profile in which a significant proportion of positive responses outweighing both negative and neutral responses. While topic 4 presents a similar emotional distribution, it has a greater degree of neutrality and a lesser degree of negativity. In general, positive prompts result in a consistent increase in positive sentiment across topics. There existed a clear match between the generative AI feedback and the intended emotional cues.

Taken together, these sentiment-topic patterns indicate that emotional prompts influence how recurrent feedback themes are framed rather than which themes appear. In all conditions, topics associated with clarity, structure, and methodological precision recur, but their sentiment profiles differ: positive prompts encapsulate these themes in praise-laden framing, neutral prompts keep them more evenly distributed while remaining skewed towards positivity, and negative prompts increase the proportion of critical evaluations while maintaining the underlying positive bias. This layer of analysis thus focuses less on reaffirming the overall positivity bias than on showing how prompt tone differently reframes otherwise stable thematic content, with alignment between intended and realized emotion strongest under positive prompts and only partial under neutral and negative prompts.

4.3 Epistemic network structures across emotionally-induced AI feedback

While sentiment and topic analyses provide an affective-thematic baseline for interpreting Epistemic Network Analysis (ENA) results, sentiment analysis indicates a consistent positivity bias across all prompt types, whereas topic modelling shows that this bias is embedded within a relatively stable set of feedback themes, including clarity of methods, organizational structure, and tone. ENA builds upon these layers by demonstrating how sentiment- and topic-informed dimensions co-occur within individual feedback episodes. In this context, the network maps are best understood as structural representations of how feedback features are connected under each prompt condition rather than as simple displays of frequency or prevalence.

ENA was used to analyse how emotional tone affects the structural organization of AI-generated feedback by modelling co-occurrence patterns between key feedback dimensions. ENA models consist of eight binary-coded attributes grouped into three functional domains: four cognitive attributes (Accuracy, Relevance, Clarity, Coherence), one dialogic attribute (Prompt_Responsiveness), and three affective attributes (Affective_Positive, Affective_Neutral, Affective_Negative). This set of attributes is represented as nodes in the ENA networks, and edges indicate the frequency and strength of their co-occurrence; thicker edges reflect stronger co-occurrence, and closer spatial proximity between nodes indicates a higher degree of conceptual clustering.

We estimated separate networks using the ENA Web Tool for positive, neutral, and negative prompt conditions. Each analysis unit corresponded to one AI-generated feedback text, and conversation boundaries were determined according to the emotional condition (i.e. tone of the prompt). The moving stanza window of 1 was selected to capture close co-occurrences

between adjacent codes while preserving the granularity of conceptual linkages within individual feedback episodes.

Figure 3a–c illustrate the resulting networks. Accuracy, Clarity, and Coherence remain structurally active for all three emotional prompt conditions. This suggests that LLMs retain basic instructional scaffolds across prompt tones. While Prompt_Responsiveness is also present in each network, its centrality and connectivity may differ. Dialogic cues are thereby prioritized differently in response to emotional framing.

The three conditions, although they share these structural anchors, differ significantly in terms of network density, node centrality, and the degree of integration between affect and cognition. The key contrast lies in how each prompt condition organizes these shared elements: positive prompts produce the most integrated architecture, neutral prompts the most fragmented one, and negative prompts a more tightly evaluative but affectively narrower structure. The positive prompt condition (Figure 3a) produces the most densely connected network, characterized by strong co-occurrence across cognitive and affective dimensions. The Affective_Positive cluster is closely associated with Clarity, Coherence and Accuracy due to its emotional affirming language embedded within pedagogically coherent feedback. Additionally, prompt responsiveness is more strongly integrated into the core network than other conditions, indicating a highly interactive and affectively attuned feedback mechanism. According to this architecture, positive prompts enable LLMs to generate mutually supportive and cognitively robust feedback.

A neutral-prompt network (Figure 3b) is the least dense and least integrated of the three networks. Affective codes (Affective_Neutral, Affective_Positive, and Affective_Negative) appear to be relatively peripheral, with a limited correlation to cognitive attributes. Co-

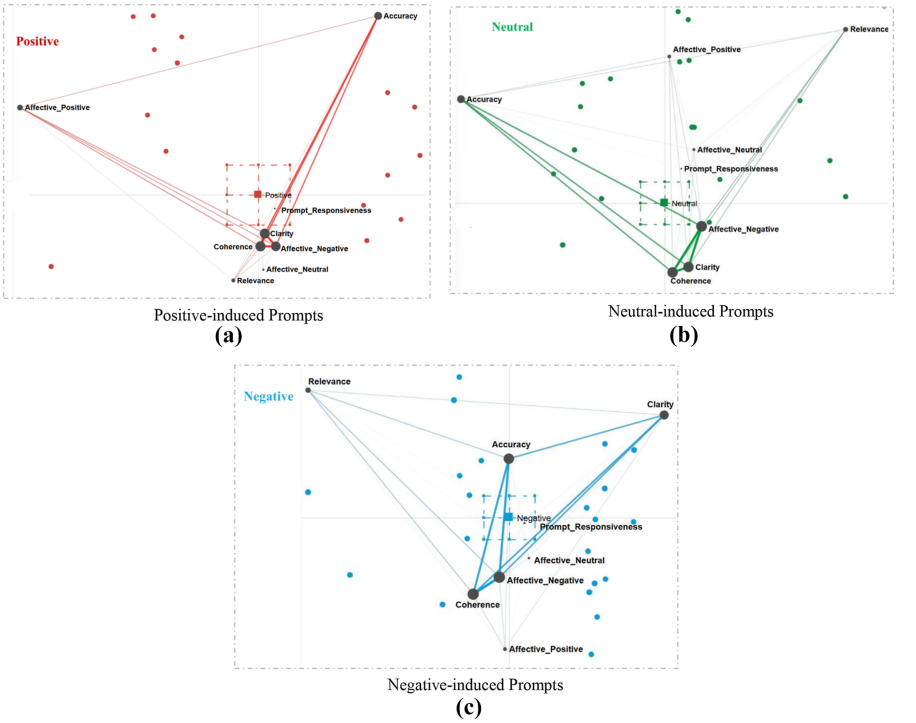

Figure 3. ENA network analysis under different emotional prompts. Authors' own work

occurrences across cognitive dimensions are also reduced to fragmented clusters rather than a unified structure. When emotional directive cues are not present, the model defaults to a more compartmentalized, less interactive mode that emphasizes factual content over dialogic or affective engagement.

In contrast, the network exhibiting a negative prompt (Figure 3c) displays a structurally polarized profile. The Affective_Negative node is prominent and closely associated with Accuracy, Prompt_Responsiveness and Clarity. The above configuration suggests that a critical tone results in evaluative, directive, and correction-oriented feedback that challenges the cognitive quality of the student's work. On the other hand, Affective_Positive and Prompt_Responsiveness are located more peripherally in the negative-prompt network, an indication of a suppression of supportive characteristics and a tendency to correct instruction. Thus, the network analysis reveals that negative emotional priming reorganizes feedback dynamics in such a way as to place precision and criticism in the forefront at the expense of interpersonal warmth and encouragement.

These ENA-based structural insights are closely aligned with earlier sentiment and topic modelling findings. Feedback generated by positive prompts was predominantly positive in sentiment and thematically coherent, in line with the interconnected and emotionally grounded structure that was observed in ENA. The neutral prompts produced the least sentiment variability and weakest alignment between topic and sentiment, as reflected in the diffuse and low-density ENA network. Despite a positive bias in sentiment analysis for negative prompts, ENA revealed a structurally distinct profile characterized by increased activation of affective negativity and stronger associations with cognitive precision. This apparent contrast suggests that while the surface emotion of feedback remains moderated by positivity, the underlying structure shifts toward greater evaluative focus and instructional rigour under negative framing. ENA thus provides further evidence that emotional prompts influence not only the sentiment distribution, but also the pedagogical architecture through which feedback is composed and delivered.

Within the Sentiment-Topic-Network framework, these ENA findings constitute the "network" layer that complements sentiment and topic analysis. The ENA models emotionally induced feedback as a pedagogical architecture characterized by specific configurations of cognitive, affective, and dialogical elements shifting with prompt tone, which has been illustrated by the contrasting networks that emerge with positive, neutral, and negative prompts. By incorporating a structural explanation of how AI designs feedback externally under different emotional framings, this integrated approach to AI-supported learning extends prior sentiment-topic approaches.

5. Discussion

We address two primary research questions in this section: the variability of AI responses to emotional triggers and the alignment of these responses with intended emotional cues. Our findings indicate that emotional prompting influences not only AI-generated feedback's polarity, but also its thematic focus and internal architecture, thereby shaping how learners may experience AI feedback as supportive, clear, and dialogic. Conceptually, the study frames emotionally induced AI feedback not only in terms of sentiment and topic, but also as a pedagogical architecture shaped by cognitive, affective, and dialogic interaction. Analyses across positive, neutral, and negative prompts highlight both a positive bias and systematic structural variation. The following subsections interpret these patterns from the perspective of feedback research and affect-aware learning design, while outlining how the integrated STN framework can inform future work linking AI feedback properties to learner outcomes via feedback architectures.

5.1 Sentiment patterns across emotionally-induced AI feedback

To answer the first research question, this study shows that generative AI tends to favour positive sentiment regardless of the emotional tone of the input prompt. A persistent positive

skew was observed. AI-generated feedback was predominantly positive, despite neutral or negative prompts. This indicates the possibility of an inherent positivity bias within the training data (Danowski *et al*., 2020; Ntoutsi *et al*., 2020; Lakkaraju *et al*., 2023; Scheel *et al*., 2021).

Negative prompts increased negative responses by a slight margin, in line with previous observations that AI responses show partial adaptation to emotional input (Srinivasan and Chander, 2021). This finding challenges earlier assumptions that AI reliably reflects prompt valence (Iqbal *et al*., 2019) or the claim that neutral prompts should naturally yield neutral emotional responses (e.g. Alasadi *et al*., 2022). In our study, neutral prompts were more likely to elicit positive responses than neutral ones. This finding has confirmed the hypothesis that AI training reflects a positivity bias.

5.2 Structural patterns in feedback configuration

Beyond sentiment-level variation, the structural composition of AI generated feedback demonstrated distinct differences based on emotional tone. Epistemic Network Analysis (ENA) shows that affective prompts alter the internal architecture by which evaluative and instructional elements are interconnected, along with emotional valence and thematic focus of each individual feedback (Hatfield, 2015).

This aligns with recent findings in AI discourse studies suggesting that feedback structure, rather than just tone, plays a critical role in shaping learner engagement and perceived support (Al Hosni, 2024; Zheng and Tse, 2023). Specifically, the co-occurrence of affective support and cognitive clarity in positively framed feedback supports previous findings showing that emotionally affirming language can enhance perceptions of feedback usefulness and coherence. Therefore, we conclude that emotional positivity contributes not only to surface sentiment but also to pedagogically integrated and dialogically responsive structures.

Comparatively, the structurally sparse and affectively detached patterns observed in the neutral-prompt condition challenge earlier assumptions regarding neutrality.

Although neutrality is often considered the default or unbiased state in instructional AI design (Iqbal *et al*., 2019), our ENA results suggest that it may instead result in disjointed and less cohesive feedback and limit its pedagogical utility (Tantry *et al.*, 2024).

Interestingly, negative prompting produced structurally tight yet affectively narrow configurations with more direct and evaluative feedback, consistent with research finding that critical tone sharpens attention to correctness while reducing perceived support (Hattie and Timperley, 2007). Under negative emotional cues, the strong structural coupling between Affective_Negative and Accuracy provides novel evidence that AI can replicate traditional teacher-centred feedback patterns (Dennis *et al*., 2018).

Overall, these findings suggest that structural feedback composition plays a latent but essential role in emotional prompt design. Specifically, sentiment and topic analysis can indicate what is expressed, while the network layer further reveals how feedback elements are organized into distinct pedagogical architectures. This study extends existing sentiment and topic-based models by showing that emotional tone determines not only what feedback contains, but also how its elements interact as well. These structural differences should be interpreted as descriptive patterns in AI-generated feedback rather than causal evidence of how emotional prompting might impact learner outcomes in an exploratory, corpus-based design. As a result, structural analyses such as ENA are highly beneficial for evaluating the coherence and interpersonal framing of AI-generated feedback. These dimensions are increasingly central to the development of emotionally intelligent educational systems (Fernandez-Nieto *et al*., 2021).

Together, the three ENA networks can be viewed as potential learning-relevant "feedback architectures" rather than just stylistic variations. As an example, positive-prompt networks approximate supportive–integrative architectures which integrate affirmation with clear and coherent task information; neutral networks approximate fragmented–minimal architectures,

which may leave learners to infer priorities on their own; negative networks approximate critical–precision architectures that sharpen accuracy while narrowing affective support. The structural perspective shows how emotional cues reorganize the internal composition of feedback-what is foregrounded, what is bundled together, and what is left peripheral-instead of treating emotional tone as a surface feature. By doing so, we provide a bridge between descriptive analysis of AI output and pedagogically meaningful questions about learners' perceptions of trust, usefulness, and guidance.

5.3 Alignment between prompt intent and AI-generated emotion

The findings indicate that positive prompts align relatively well with intended emotional cues, whereas negative and neutral prompts cause this alignment to weaken considerably. The persistence of positive sentiment in response to neutral and negative cues demonstrates the challenges associated with achieving true emotional neutrality within AI-generated feedback (Ballotta *et al*., 2023). Even when given negative prompts, generative AI often defaults to neutral or mildly positive tones. The partial alignment is also evident in Schmidt *et al*. (2020)'s study, where generative AI failed to fully adapt to negative emotional cues. As AI-generated feedback barely captures and adapts to negative emotional cues, there might be an overrepresentation of positive examples in the training data or inherent algorithmic biases towards positive sentiment (Ferrara, 2023) that contribute to this tendency.

Moreover, the study observed significant variations in emotional intensity across different topics, even when prompted with the same type of question. The result implied that the thematic elements play a crucial role in AI's sentiment generation. Thematic content might influence how AI-generated feedback expresses emotion, as certain topics consistently elicit stronger responses. This observation aligns with findings by Hartung *et al*. (2023), who emphasize the importance of thematic context in shaping AI-generated sentiment. A similar study conducted by Broekens *et al*. (2023) also demonstrated that thematic cues can aid in eliciting more nuanced sentiment analyses and emotions in large language models.

This study demonstrates how contextual factors such as theme content and prompt design shape emotional variability in AI-generated text. By calibrating prompts and refining training data, designers may achieve closer alignment between intended emotional cues and emotional interpretations in AI feedback. Together, these insights provide a theoretical and practical foundation for future research on how prompt-emotion alignment affects learners' experiences with AI-supported feedback.

5.4 Implications for learning and feedback design

While the current study did not measure learning outcomes directly, the combined sentiment, topic, and ENA results suggest that emotional prompts significantly alter the way cognitive and affective elements are incorporated into feedback structures that learners encounter, in addition to simply colouring AI feedback. Consequently, this has implications for how learners may experience, interpret, and use AI-generated feedback in instructional contexts, and for how instructors and designers can incorporate emotional cues into prompt templates. Viewed through established learning perspectives, these differences matter because feedback supports learning not only by conveying information, but also by shaping the interpretive, motivational, and relational conditions under which learners engage with it (Carless and Boud, 2018; Winstone *et al*., 2017). The three patterns result in three empirically based feedback architectures - supportive-integrative, fragmented-minimal, and critical-precision - that form a compact, data-driven model.

On this basis, we conceptualize an emotionally intelligent AI tutor as a system that can select and sequence feedback architectures in response to the learner's state and the task at hand. Minimal adaptive models (1) emphasize supportive–integrative feedback when learners are drafting in early stages or indicate low confidence, (2) provide brief instances of critical–precision feedback when a draft is reasonably coherent and learners are ready to focus on their

errors, and (3) reserve fragmented-minimal feedback for diagnostic or grading-related tasks that require minimal affect. The model directly translates our empirical network patterns into operational design rules that can be implemented through prompt templates, interface options, or higher-level orchestration in AI-driven tutoring applications.

From a theoretical perspective, these findings contribute to research on feedback and affect-aware learning design in three ways. First, they operationalize the broad idea of "emotionally intelligent" feedback into three empirically grounded architectures of supportive-integrative, fragmented-minimal, and critical-precision, which can be analysed at the level of sentiment, thematic focus, and structural coupling between affective and cognitive aspects. Second, the study also offers a multilayered view of how emotional prompts reconfigure feedback by combining sentiment analysis, topic modelling, and ENA within a single Sentiment-Topic-Network framework, moving beyond prior work that has largely focused on surface valence changes as the only way to understand emotion. Third, the study specifies a tractable design space in which future learning research can generate and test hypotheses about revision behaviour, epistemic trust, and emotional safety in AI-supported feedback by linking these architectures to the adaptive model shown above. In this sense, the STN framework functions both as a diagnostic tool for analysing AI-generated feedback and as a theoretical foundation for developing emotionally tuned AI tutors.

Our data also reveal additional practical implications for learners. First, persistent positivity bias across all three prompt conditions suggest that generative AI is more likely to produce feedback that sounds encouraging rather than feedback that fully reflects the intended emotional tone. Despite neutral and negative prompts, a majority of feedback segments were classified as positive, and neutral prompts especially resulted in unexpectedly positive sentiment distributions. Based on this pattern, learners may encounter feedback that is perceived as supportive, even when teachers or designers intend their evaluation to be neutral. This bias can have a double-edged effect from a learning perspective: while positive tone has been associated with perceived support and willingness to engage with feedback, it may also obscure the seriousness of problems or dilute the sense of urgency surrounding revision.

Secondly, the ENA results indicate that positive prompts produce the most densely connected networks, where affective support such as praise is closely related to clarity, coherence, and accuracy. Emotional affirmation results in feedback both encouraging and pedagogically coherent. Due to prior feedback research linking clear, coherent explanations and references to learners' work to stronger engagement and perceived usefulness, this configuration is pedagogically relevant. Positively framed prompts can increase engagement and perceived usefulness among learners by embedding cognitive guidance within a supportive interpersonal framework. This will make it easier for learners to accept and act upon critique.

Third, the neutral prompt condition suggests that emotional neutrality is not a benign default. Alternatively, the neutral-prompt network exhibited structural sparsity and emotional disengagement, which mirrored the thematically shallow, positive-skewed results observed in the topic and sentiment analyses. This, coupled with the sentiment analysis, suggests a "hin" feedback architecture: comments that sound polite and moderately positive, but are less structurally integrated. Such feedback may be perceived as vague or generic, providing neither strong emotional support nor tightly linked cognitive guidance for learning. As such, it challenges the assumption that neutral prompting is automatically the safest or most objective choice when using AI-mediated feedback.

Fourth, negative prompts, despite their thematic richness and greater affective diversity, revealed feedback centred around Affective_Negative and Accuracy, reflecting a more evaluative and corrective tone. ENA clarified that emotional tone does not only influence what is said, but also how concepts are clustered, which reveals the latent differences between instructional feedback architectures. This configuration may be beneficial from a learning perspective because students receive targeted information about what is wrong and why, and it may also be beneficial for diagnostic precision. However, this network's relative peripheral

position suggests a potential trade-off: feedback may be cognitively beneficial, but is less likely to be experienced as motivating or supportive, especially by learners who are less confident or less familiar with the task.

Across conditions, the STN framework used in this study indicates that learning-relevant properties of feedback are emergent: they are determined both by the content of individual sentences and by the way affective and cognitive features are linked across the feedback message. Rather than treating AI feedback as a monolithic product, our findings show that the same student text can produce qualitatively different feedback structures solely as a result of the prompt tone. Thus, emotional prompt design is not a cosmetic choice but rather an element of learning design that affects learners' understanding, value, and response to feedback. These structural insights demonstrate how emotional cues implicitly guide the assembly of affective and cognitive features in AI communication and bridge descriptive analysis and instructional decision-making.

5.5 Design guidelines for emotion-aware AI feedback in higher education

We outline several design guidelines for using emotion-laden prompts in generative AI to support feedback in higher education settings similar to postgraduate scientific writing and supervisor-student feedback workflows. From a practical standpoint, these guidelines represent the main implications of the study for classroom teaching, supervisor-student feedback practices, and the design of educational AI tools. Several studies have indicated that generative AI can be more effective at engaging learners when it tailors emotionally resonant content, and that AI-mediated instruction should align thematic content and emotional tone (Dickey and Bejarano, 2023; Keshishi and Hack, 2023; Kadaruddin, 2023). This work contributes to that line of research by showing how different emotional cues reorganize feedback architectures in ways that can be intentionally leveraged in course and assessment design. This practical relevance matters particularly in higher education contexts, where feedback design must support not only informational clarity, but also learner confidence, engagement, and sustained revision. We therefore limit our claims to these immediate educational settings and do not infer broader societal or public policy effects from a single dataset.

To begin with, align the tone of the prompt with the pedagogical goal of the feedback episode. Our findings suggest that carefully crafted positive prompts are likely to result in feedback that combines affective support and clear, coherent explanations when the primary goal is to sustain engagement, normalize difficulty, or reduce anxiety around complex tasks. A well-designed prompt can, for instance, be used to create learning materials that are adaptable to the emotional states of students, thereby promoting motivation and improving retention of information. Instructors should specify in the prompt that AI highlight concrete suggestions for improvement and promote the integrated affective-cognitive structures observed in the positive-prompt ENA networks. Emotional alignment may be incorporated into AI-generated content in order to meet diverse learning needs and create environments that are more personalized, impactful, and personal (Dickey and Bejarano, 2023).

Secondly, neutral prompts should not be treated as a default for "objective" feedback. Neutral condition produced the least structurally integrated feedback, even though the surface sentiment was moderately positive. By defaulting to neutral phrasing, designers may unintentionally generate feedback which feels polite but shallow. For example, when instructors wish to limit emotional colouring in summative contexts, prompts may need to explicitly specify structurally elaborated explanations (e.g. "provide specific, step-by-step suggestions linked to the student's text") rather than relying solely on neutral tone.

Thirdly, use critical or negative prompts selectively and in conjunction with buffering strategies. Negative prompts produced networks in which Affective_Negative was tightly coupled with Accuracy and Prompt_Responsiveness. This indicates that they are effective when precise diagnosis and error correction are required. The reduced presence of affective

support in these networks may make such prompts more appropriate when learners are already relatively confident, or when critical feedback is embedded within a broader sequence that includes more supportive messages. In practice, a sequential prompting strategy may be employed: first eliciting a positive, supportive overview, followed by a more critical, accuracy-focused feedback round, so that learners are provided with both reassurance and clarity on their next steps.

Fourth, explicitly define the role of AI feedback within the wider feedback ecology. Considering that emotional prompts have the potential to substantially alter feedback structure, learners may benefit from explicit explanations that AI comments are shaped by prompt design and are intended to complement rather than replace human judgement. For example, instructors could disclose that "this feedback was generated with a prompt emphasizing encouragement and clarity," or conversely, "this feedback was generated with a prompt emphasizing the need to be particularly critical about methodological details." Transparency in feedback may assist students in interpreting the tone and focus more accurately and integrating it with teacher feedback and peer feedback.

Lastly, consider the emotional prompt design in learning research and course design as a variable, rather than just a detail on the interface. As this study demonstrates that prompt tone significantly impacts sentiment distributions, topic-sentiment alignments, and network structures, future learning design may also evaluate the effects of different emotional prompting strategies across tasks, cohorts, or disciplines as well as downstream indicators such as revision quality, help-seeking behaviour, and perception of trust in artificial intelligence-supported feedback across tasks, cohorts, or disciplines.

In practice, educators and developers can immediately take advantage of this framework in three specific ways. Firstly, instructors can embed prompt templates corresponding to each feedback architecture (supportive-integrative, fragmented-minimal, critical-precision) into local course guidelines and instruct students as to when and how they should use them. In addition, developers of institutional AI tools should offer a feedback mode selector that maps directly onto these architectures, to allow teachers to align system behaviour with course-level pedagogy rather than accepting a single default setting. Third, programme leaders may develop brief educational materials aimed at learners and teachers describing the affordances and risks of each mode, in order to assist them in interpreting AI feedback appropriately and combining it with human feedback. More broadly, a variety of contextually appropriate emotional design elements, including empathic tutor messages, affectively framed narratives, and supportive feedback, have been found to promote engagement, understanding, and affective or socio-emotional outcomes (D'Mello and Graesser, 2012; Lehman, D'Mello and Graesser, 2012; Um *et al*., 2012; Pekrun *et al*., 2011; Artino and Jones, 2012).

Over all, this work contributes to learning and instructional design by showing empirically that emotional prompts modify the architecture of AI feedback in ways that have pedagogical implications. This implications section provides concrete design suggestions based on observed differences between positive, neutral, and negative prompt conditions rather than offering general or speculative recommendations. This provides a tangible bridge between theory and practice in AI-mediated feedback design. Thus, we restrict these implications to societal and public policy implications at the institutional and classroom levels in higher education and do not infer system-wide impacts from this single dataset.

Our research implications are based on three hypotheses that could be empirically tested in the future. Together, they translate our descriptive findings into testable predictions about learning outcomes:

H1. As compared to fragmented–minimal feedback, supportive–integrative feedback will be associated with a higher level of perceived usefulness, emotional support, and revision self-efficacy.

H2. For advanced learners and late drafting stages, critical-precision feedback will be associated with more substantive revisions than supportive-integrative feedback, but will also be associated with a lower level of perceived trust.

H3. Sequencing supportive-integrative feedback with brief episodes of critical-precision feedback will lead to higher revision quality and equally high or higher perceived trust than relying exclusively on either architecture.

6. Conclusion

The present study examined how a large language model generates feedback in response to emotion-laden prompts, integrating sentiment analysis, topic modelling, and Epistemic Network Analysis to characterize both the surface tone and internal structure of AI feedback. A robust positivity bias was observed across 198 feedback messages concerning postgraduate scientific reports: positive sentiment dominated even when prompts were neutral or explicitly negative, and only the negative condition resulted in a substantial increase in the proportion of negative codes. This topic-sentiment pattern was also found to be influenced by thematic content, with some topics consistently attracting more intense emotional expressions than others. This study indicates that emotional prompting is indeed associated with the affective profile of AI feedback, as indicated in the first research question. Regarding the second question, this adaptation is only partial, with feedback's valence only imperfectly reflecting the emotional tone of prompts.

ENA extends this description from "what is said" to "how feedback is assembled". Three distinct feedback architectures were observed under the three emotional conditions: supportive–integrative networks under positive prompts that were closely related to affect and cognition; fragmented–minimal networks under neutral prompts with weaker links between affect and cognition; and critical–precision networks under negative prompts that emphasized accuracy and directive comments and suppressed positive affect. As a result, these profiles offer a data-driven vocabulary and set of visualisations that learning scientists and designers can use to determine the likely experiential benefits of different AI feedback configurations, even prior to the assessment of learner outcomes.

A major contribution of our work to learning research has been to characterize feedback architectures and designable regimes rather than to demonstrate direct improvements in learning outcomes. The sentiment, topic, and network analyses identify empirically grounded patterns of feedback learners would receive if instructors or systems adopted different prompting strategies. Supportive-integrative, fragmented-minimal, and critical-precision feedback architectures transform network diagrams into testable hypotheses about learner experience and motivation. However, we did not measure achievement gains, long-term retention, or metacognitive skill development. The implications we draw for feedback design and tutoring systems should be understood as theoretically based and empirically based, but still requiring prospective evaluation in classrooms and assessment settings. We have therefore mapped the design space of emotionally tuned AI feedback, and proposed methods for connecting these configurations to concrete learning processes and outcomes, without overstating causal effects that our study was not designed to assess.

The study is limited by its use of a limited, context-bound dataset. The emotional prompts were predefined and stylized to represent three broad supervisory tones (positive, neutral, negative), which cannot fully capture the complexity and nuance of real-world emotional expressions. In addition, the data were collected from only 66 postgraduate students at a single institution, resulting in a relatively small and non-representative sample. Our findings, therefore, should be interpreted as exploratory patterns rather than statistically generalizable estimates of how all generative AI systems respond to emotional cues. Also, the study heavily relies on quantitative sentiment analysis, which, while systematic, may underrepresent the richness of emotional expression that can be captured by qualitative approaches.

Future research should consider more complex emotional contexts and prompt designs, as well as employ mixed-method approaches to assess AI's response to a broad range of emotional cues, including human assessment of structural coherence, emotional authenticity, and subtle affective signals. Incorporating qualitative analyses alongside quantitative sentiment and network metrics will enable researchers to capture nuances of AI-generated emotional expression beyond simple classifications based on polarity. Such work will help refine the Sentiment-Topic-Network framework and advance the alignment of generative AI with human emotional expression in educational and other high-stakes settings.

Declaration

All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Faculty of Applied Sciences, Macao Polytech University (RP/FCA-08/2023).

References

Abdullah, H.M.A., Malkana, G.A., Javed, M. and Shahzaib, M.M. (2024), "Exploring the ethical implications of ChatGPT in medical education: privacy, accuracy, and professional integrity in a cross-sectional survey", *Cureus*, Vol. 16 No. 12, e75895, doi: 10.7759/cureus.75895.

Adie, L., van der Kleij, F. and Cumming, J. (2018), "The development and application of coding frameworks to explore dialogic feedback interactions and self-regulated learning", *British Educational Research Journal*, Vol. 44 No. 4, pp. 704-723, doi: 10.1002/berj.3463.

Ahmed, R. (2024), "Exploring ChatGPT usage in higher education: patterns, perceptions,and ethical implications among university students", *Journal of Digital Learning and Distance Education*, Vol. 3 No. 6, pp. 1122-1131, doi: 10.56778/jdlde.v3i6.363.

Aini, L.R., Nurfadhilah, E., Jarin, A., Santosa, A. and Uliniansyah, M.T. (2023), "Enhancing sentiment analysis models through multi-technique data augmentation: a study with IndoBERT", *2023 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*, pp. 137-142, doi: 10.1109/IC3INA60834.2023.10285775.

Al Hosni, J. (2024), "Stylometric analysis of AI chatbot-generated emails: are students losing their linguistic fingerprint?", *Journal of English Language Teaching and Applied Linguistics*, Vol. 6 No. 3, pp. 33-42, doi: 10.32996/jeltal.2024.6.3.5.

Alasadi, J., Hilli, A., Atrey, P. and Singh, V. (2022), "A generative approach to mitigate bias in face matching using learned latent structure", *2022 IEEE Eighth International Conference on Multimedia Big Data (BigMM)*, pp. 150-157, doi: 10.1109/BigMM55396.2022.00033.

Artino, A.R. Jr and Jones, K.D., II (2012), "Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning", *The internet and higher education*, Vol. 15 No. 3, pp. 170-175, doi: 10.1016/j.iheduc.2012.01.006.

Ba, S., Hu, X., Stein, D. and Liu, Q. (2024), "Anatomizing online collaborative inquiry using directional epistemic network analysis and trajectory tracking", *British Journal of Educational Technology*, Vol. 55 No. 5, pp. 2173-2191, doi: 10.1111/bjet.13441.

Ballotta, D., Maramotti, R., Borelli, E., Lui, F. and Pagnoni, G. (2023), "Neural correlates of emotional valence for faces and words", *Frontiers in Psychology*, Vol. 14, 1055054, doi: 10.3389/fpsyg.2023.1055054.

Banihashem, S.K., Noroozi, O., Van Ginkel, S., Macfadyen, L.P. and Biemans, H.J. (2022), "A systematic review of the role of learning analytics in enhancing feedback practices in higher education", *Educational Research Review*, Vol. 37, 100489, doi: 10.1016/j.edurev.2022.100489.

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003), "Latent dirichlet allocation", *Journal of Machine Learning Research*, Vol. 3, Jan, pp. 993-1022, doi: 10.1162/jmlr.2003.3.4-5.993.

Broekens, J., Hilpert, B., Verberne, S., Baraka, K., Gebhard, P. and Plaat, A. (2023), "Fine-grained affective processing capabilities emerging from large language models", *2023 11th International*

Conference on Affective Computing and Intelligent Interaction (ACII), pp. 1-8, doi: 10.1109/ACII59096.2023.10388177.

Carless, D. and Boud, D. (2018), "The development of student feedback literacy: enabling uptake of feedback", *Assessment and Evaluation in Higher Education*, Vol. 43 No. 8, pp. 1315-1325, doi: 10.1080/02602938.2018.1463354.

Chiang, C.-W., Lu, Z., Li, Z. and Yin, M. (2024), "Enhancing AI-assisted group decision making through LLM-powered devil's advocate", *Proceedings of the 29th International Conference on Intelligent User Interfaces (IUI '24)*, pp. 103-119, doi: 10.1145/3640543.3645199.

Cowling, M., Crawford, J., Allen, K.A. and Wehmeyer, M. (2023), "Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision", *Australasian Journal of Educational Technology*, Vol. 39 No. 4, pp. 89-103, doi: 10.14742/ajet.8598.

Danowski, J., Yan, B. and Riopelle, K. (2020), "A semantic network approach to measuring sentiment", *Quality and Quantity*, Vol. 55 No. 1, pp. 221-255, doi: 10.1007/s11135-020-01000-x.

Dennis, A.A., Foy, M.J., Monrouxe, L.V. and Rees, C.E. (2018), "Exploring trainer and trainee emotional talk in narratives about workplace-based feedback processes", *Advances in Health Sciences Education*, Vol. 23 No. 1, pp. 75-93, doi: 10.1007/s10459-017-9775-0.

Dickey, E. and Bejarano, A. (2023), "A model for integrating generative AI into course content development", arXiv e-prints, arXiv-2308, doi: 10.48550/2308.12276.

D'Mello, S. and Graesser, A. (2012), "Dynamics of affective states during complex learning", *Learning and Instruction*, Vol. 22 No. 2, pp. 145-157, doi: 10.1016/j.learninstruc.2011.10.001.

Fatima-Zahrae, S., Wafae, S. and Amal, E.M. (2021), "Application of latent dirichlet allocation (LDA) for clustering financial tweets", *E3S Web of Conferences*, Vol. 297, 01071, EDP Sciences, doi: 10.1051/e3sconf/202129701071.

Fernandez-Nieto, G.M., Martinez-Maldonado, R., Kitto, K. and Buckingham Shum, S. (2021), "Modelling spatial behaviours in clinical team simulations using epistemic network analysis: methodology and teacher evaluation", *LAK2111th International Learning Analytics and Knowledge Conference*, pp. 386-396, doi: 10.1145/3448139.3448176.

Ferrara, E. (2023), "Should ChatGPT be biased? Challenges and risks of bias in large language models", ArXiv,abs/2304.03738, doi: 10.48550/arXiv.2304.03738.

Ficler, J. and Goldberg, Y. (2017), "Controlling linguistic style aspects in neural language generation", arXiv preprint arXiv:1707.02633, doi: 10.48550/arXiv.1707.02633.

Gilardi, F., Alizadeh, M. and Kubli, M. (2023), "ChatGPT outperforms crowd workers for text-annotation tasks", *Proceedings of the National Academy of Sciences*, Vol. 120 No. 30, e2305016120, doi: 10.1073/pnas.2305016120.

Hartung, K., Herygers, A., Kurlekar, S., Zakaria, K., Volkan, T., Gröttrup, S. and Georges, M. (2023), "Measuring sentiment bias in machine translation", *International Conference on Text, Speech and Dialogue*, pp. 82-93, doi: 10.1007/978-3-031-40498-6_8.

Hatfield, D. (2015), "The right kind of telling: an analysis of feedback and learning in a journalism epistemic game", *International Journal of Gaming and Computer-Mediated Simulations*, Vol. 7 No. 2, pp. 1-23, doi: 10.4018/IJGCMS.2015040101.

Hattie, J. and Timperley, H. (2007), *The power of feedback. Review of Educational Re-search*, Vol. 77 No. 1, pp. 81-112, doi: 10.3102/003465430298487.

Herb, A. and Lloyd, C. (2024), "The future of feedback: exploring the use of generative AI", in *Formative Assessment*, ASCILITE Publications, pp. 141-142, doi: 10.14742/apubs.2024.1386.

Höpken, W., Fuchs, M., Menner, T. and Lexhagen, M. (2017), "Sensing the online social sphere using a sentiment analytical approach", in Xiang, Z. and Fesenmaier, R. (Eds), *Analytics in Smart Tourism Design: Concepts and Methods*, Springer, pp. 129-146, doi: 10.1007/978-3-319-44263-1_8.

Hu, Z., Yang, Z., Liang, X., Salakhutdinov, R. and Xing, E.P. (2017), "Toward controlled generation of text", *Proceedings of the 34th International Conference on Machine Learning*, Vol. 70, PMLR, pp. 1587-1596, doi: 10.48550/arXiv.1703.00955.

Iqbal, M., Karim, A. and Kamiran, F. (2019), "Balancing prediction errors for robust sentiment classification", *ACM Transactions on Knowledge Discovery from Data*, Vol. 13 No. 3, pp. 1-21, doi: 10.1145/3328795.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y.J., Madotto, A. and Fung, P. (2023), "Survey of hallucination in natural language generation", *ACM Computing Surveys*, Vol. 55 No. 12, pp. 1-38, doi: 10.1145/3571730.

Kadaruddin, K. (2023), "Empowering education through Generative AI: innovative instructional strategies for tomorrow's learners", *International Journal of Business, Law, and Education*, Vol. 4 No. 2, pp. 618-625, doi: 10.56442/ijble.v4i2.215.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Kasneci, G., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J. and Kuhn, J. (2023), "ChatGPT for good? On opportunities and challenges of large language models for education", *Learning and Individual Differences*, Vol. 103, 102274, doi: 10.1016/j.lindif.2023.102274.

Keshishi, N. and Hack, S. (2023), "Emotional intelligence in the digital age: harnessing AI for students' inner development", *Journal of Perspectives in Applied Academic Practice*, Vol. 11 No. 3, pp. 172-175, doi: 10.56433/jpaap.v11i3.579.

Keskar, N.S., McCann, B., Varshney, L.R., Xiong, C. and Socher, R. (2019), "Ctrl: a conditional transformer language model for controllable generation", arXiv preprint arXiv:1909.05858, doi: 10.48550/arXiv.1909.05858.

Kit, Y. and Mokji, M. (2022), "Sentiment analysis using pre-trained language model with no fine-tuning and less resource", *IEEE Access*, Vol. 10, pp. 107056-107065, doi: 10.1109/ACCESS.2022.3212367.

Korinek, A. (2023), "Generative AI for economic research: use cases and implications for economists", *Journal of Economic Literature*, Vol. 61 No. 4, pp. 1281-1317, doi: 10.1257/jel.20231736.

Kumawat, S., Yadav, I., Pahal, N. and Goel, D. (2021), "Sentiment analysis using language models: a study", *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 984-988, doi: 10.1109/Confluence51648.2021.9377043.

Kuzminykh, I., Nawaz, T., Shenzhang, S., Ghita, B., Raphael, J. and Xiao, H. (2024), "Personalised feedback framework for online education programmes using generative AI", arXiv preprint arXiv:2410.11904, doi: 10.48550/arXiv.2410.11904.

Lakkaraju, K., Srivastava, B. and Valtorta, M.G. (2023), "Rating sentiment analysis systems for bias through a causal lens", *IEEE Transactions on Technology and Society*, Vol. 5 No. 1, pp. 82-92, doi: 10.1109/TTS.2024.3375519.

Lee, S. and Park, G. (2023), "Exploring the impact of ChatGPT literacy on user satisfaction: the mediating role of user motivations", *Cyberpsychology, Behavior, and Social Networking*, Vol. 26 No. 12, pp. 913-918, doi: 10.1089/cyber.2023.0312.

Lehman, B., D'Mello, S. and Graesser, A. (2012), "Confusion and complex learning during interactions with computer learning environments", *The Internet and Higher Education*, Vol. 15 No. 3, pp. 184-194, doi: 10.1016/j.iheduc.2012.01.002.

Li, C., Wang, J., Zhu, K., Zhang, Y., Hou, W., Lian, J. and Xie, X. (2023), "EmotionPrompt: leveraging psychology for large language models enhancement via emotional stimulus", abs/2307.11760, doi: 10.48550/arXiv.2307.11760.

Lin, C., Ibeke, E., Wyner, A. and Guerin, F. (2015), "Sentiment–topic modelling in text mining", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 5 No. 5, pp. 246-254, doi: 10.1002/widm.1161.

Mao, R., Liu, Q., He, K., Li, W. and Cambria, E. (2023), "The biases of pre-trained language models: an empirical study on prompt-based sentiment analysis and emotion detection", *IEEE Transactions on Affective Computing*, Vol. 14 No. 3, pp. 1743-1753, doi: 10.1109/TAFFC.2022.3204972.

Naz, I. and Robertson, R. (2024), "Exploring the feasibility and efficacy of ChatGPT3 for personalized feedback in teaching", *Electronic Journal of E-Learning*, Vol. 22 No. 2, pp. 98-111, doi: 10.34190/ejel.22.2.3345.

Nguyen, T.T., Hatua, A. and Sung, A.H. (2023), "How to detect AI-generated texts?", *IEEE Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 0464-0471, doi: 10.1109/UEMCON59035.2023.10316132.

Nikolopoulou, K. (2024), "Generative artificial intelligence in higher education: exploring ways of harnessing pedagogical practices with the assistance of ChatGPT", *International Journal of Changes in Education*, Vol. 1 No. 2, pp. 103-111, doi: 10.47852/bonviewijce42022489, available at: https://orcid.org/0000-0002-2175-1765

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., Broelemann, K., Kasneci, G., Tiropanis, T. and Staab, S. (2020), "Bias in data-driven artificial intelligence systems—an introductory survey", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 10 No. 3, e1356, doi: 10.1002/widm.1356.

Ocampo, T.S.C., Silva, T.P., Alencar-Palha, C., Haiter-Neto, F. and Oliveira, M.L. (2023), "ChatGPT and scientific writing: a reflection on the ethical boundaries", *Imaging science in dentistry*, Vol. 53 No. 2, p. 175, doi: 10.5624/isd.20230085.

Ohmura, M., Kakusho, K. and Okadome, T. (2014), "Tweet sentiment analysis with latent dirichlet allocation", *International Journal of Information Retrieval Research*, Vol. 4 No. 3, pp. 66-79, doi: 10.4018/IJIRR.2014070105.

Osmani, A., Mohasefi, J.B. and Gharehchopogh, F.S. (2020), "Enriched latent dirichlet allocation for sentiment analysis", *Expert Systems*, Vol. 37 No. 4, e12527, doi: 10.1111/exsy.12527.

Ouyang, L., Wu, J., Jiang, C., Almeida, D., Wainwright, C., Mishkin, P. and Leike, J. (2022), "Training language models to follow instructions with human feedback", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, pp. 2305-2316, doi: 10.48550/arXiv.2203.02155.

Pang, T.Y., Kootsookos, A. and Cheng, C.T. (2024), "Artificial intelligence use in feedback: a qualitative analysis", *Journal of University Teaching and Learning Practice*, Vol. 21 No. 6, pp. 108-125, doi: 10.53761/40wmcj98, available at: https://search.informit.org/doi/10.3316/informit.T202409290000369164832 3996

Pekrun, R., Goetz, T., Frenzel, A.C., Barchfeld, P. and Perry, R.P. (2011), "Measuring emotions in students' learning and performance: the Achievement Emotions Questionnaire (AEQ)", *Contemporary Educational Psychology*, Vol. 36 No. 1, pp. 36-48, doi: 10.1016/j.cedpsych.2010.10.002.

Rashkin, H., Smith, E.M., Li, M. and Boureau, Y.L. (2019), "Towards empathetic open-domain conversation models: a new benchmark and dataset", *Proceed-ings of the 57th annual meeting of the association for computational linguistics*, pp. 5370-5381, doi: 10.18653/v1/P19-1534.

Röder, M., Both, A. and Hinneburg, A. (2015), "Exploring the space of topic coherence measures", *Proceedings of the eighth ACM international conference on Web search and data mining*, pp. 399-408, doi: 10.1145/2684822.268532.

Ruotsalo, T., Mäkelä, K., Spapé, M.M.A. and Leiva, L.A. (2023), "Affective relevance: inferring emotional responses via fNIRS neuroimaging", *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1796-1800, doi: 10.1145/3539618.3591946.

Scheel, A.M., Schijen, M.R. and Lakens, D. (2021), "An excess of positive results: comparing the standard psychology literature with registered reports", *Advances in Methods and Practices in Psychological Science*, Vol. 4 No. 2, doi: 10.1177/25152459211007467.

Schmidt, S., Sojer, C., Hass, J., Kirsch, P. and Mier, D. (2020), "fMRI adaptation reveals: the human mirror neuron system discriminates emotional valence", *Cortex*, Vol. 128, pp. 270-280, doi: 10.1016/j.cortex.2020.03.026.

See, A., Roller, S., Kiela, D. and Weston, J. (2019), "What makes a good conversation? How controllable attributes affect human judgments", arXiv preprint arXiv:1902.08654, doi: 10.48550/arXiv.1902.08654.

Shute, V.J. (2008), "Focus on formative feedback", *Review of Educational Research*, Vol. 78 No. 1, pp. 153-189, doi: 10.3102/0034654307313795.

Smit, R., Hess, K., Taras, A., Bachmann, P. and Dober, H. (2023), "The role of interactive dialogue in students' learning of mathematical reasoning: a quantitative multi-method analysis of feedback episodes", *Learning and Instruction*, Vol. 86, 101777, doi: 10.1016/j.learninstruc.2023.101777.

Srinivasan, R. and Chander, A. (2021), "Biases in AI systems", *ACM Queue*, Vol. 19 No. 2, pp. 45-64, doi: 10.1145/3466132.3466134.

Sun, X., Li, X., Zhang, S., Wang, S., Wu, F., Li, J., Zhang, T. and Wang, G. (2023), "Sentiment analysis through LLM negotiations", abs/2311.01876, doi: 10.48550/arXiv.2311.01876.

Taj, N. and Girisha, G.S. (2024), "Leveraging natural language processing for in-depth analysis and insights from hospital patient feedback data", *2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, IEEE, pp. 01-08, doi: 10.1109/ICDCECE60827.2024.10548960.

Tantry, R., Shenoy, S.U., Acharya, S. and Prathibha, K.N. (2024), "Artificial intelligence assisted student states monitoring based on enhancing cognitive and emotional feedback mechanism using collaborative learning environments", *2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, IEEE, pp. 1-5, doi: 10.1109/ACCAI61061.2024.10601800.

Telmo, K.A.B., Alviola, K.V., Desamparado, J.J.S., Cabigan, J.N.A., Santiago, C.S. Jr and Shimada, R.A.A. (2024), "Sentiment analysis of student evaluation for teachers using valence-aware dictionary and sentiment reasoner", *Journal of Interdisciplinary Perspectives*, Vol. 2 No. 9, p. 1, doi: 10.69569/jip.2024.0328.

Um, E., Plass, J.L., Hayward, E.O. and Homer, B.D. (2012), "Emotional design in multimedia learning", *Journal of Educational Psychology*, Vol. 104 No. 2, pp. 485-498, doi: 10.1037/a0026609.

Viberg, O., Baars, M., Mello, R.F., Weerheim, N., Spikol, D., Bogdan, C., Gasevic, D. and Paas, F. (2024), "Exploring the nature of peer feedback: an epistemic network analysis approach", *Journal of Computer Assisted Learning*, Vol. 40 No. 6, pp. 2809-2821, doi: 10.1111/jcal.13035.

Vickneswaran, J., Navanesan, P., Vijayaratnam, V. and Thayasivam, U. (2020), "Simplified approach for predicting emotions of multi-turn textual utterances", *2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer)*, pp. 71-76, doi: 10.1109/ICTer51097.2020.9325442.

Williams, R.T. (2024), "The ethical implications of using generative chatbots in higher education", *Frontiers in Education*, Vol. 8, 1331607, doi: 10.3389/feduc.2023.1331607.

Winstone, N.E., Nash, R.A., Parker, M. and Rowntree, J. (2017), "Supporting learners' agentic engagement with feedback: a systematic review and a taxonomy of recipience processes", *Educational Psychologist*, Vol. 52 No. 1, pp. 17-37, doi: 10.1080/00461520.2016.1207538.

Xu, P. and Jin, Y. (2020), "Coherence and stochastic resonance in a second-order asymmetric tri-stable system with memory effects", *Chaos, Solitons and Fractals*, Vol. 138, 109857, doi: 10.1016/j.chaos.2020.109857.

Yang, T., Yao, R., Yin, Q., Tian, Q. and Wu, O. (2020), "Mitigating sentimental bias via a polar attention mechanism", *International Journal of Data Science and Analytics*, Vol. 11 No. 1, pp. 27-36, doi: 10.1007/s41060-020-00231-3.

Zhang, T., Irsan, I., Thung, F. and Lo, D. (2023), "Revisiting sentiment analysis for software engineering in the era of large language models", *Proceedings of the International Conference on Software Engineering*, Vol. 5 No. 3, pp. 150-161, doi: 10.1145/3697009.

Zhao, Y., Huang, Z., Seligman, M. and Peng, K. (2024), "Risk and prosocial behavioural cues elicit human-like response patterns from AI chatbots", *Scientific Reports*, Vol. 14 No. 1, p. 7095, doi: 10.1038/s41598-024-55949-y.

Zheng, W. and Tse, A.W.C. (2023), "The impact of generative artificial intelligence-based formative feedback on the mathematical motivation of Chinese grade 4 students: a case study", *2023 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*, IEEE, pp. 1-8, doi: 10.1109/TALE56641.2023.10398319.

Zhou, H., Huang, M., Zhang, T., Zhu, X. and Liu, B. (2018), "Emotional chatting machine: emotional conversation generation with internal and external memory", *Proceedings of the AAAI conference on artificial intelligence*, Vol. 32 No. 1, doi: 10.1609/aaai.v32i1.11325.

Zhu, L., Xu, Y., Zhu, Z., Bao, Y. and Kong, X. (2022), "Fine-grained sentiment-controlled text generation approach based on pre-trained language model", *Applied Sciences*, Vol. 13 No. 1, p. 264, doi: 10.3390/app13010264.

Further reading

Fu, X.-Y., Laskar, Md.T.R., Chen, C. and Shashi Bhushan, Tn (2023), "Are large language models reliable judges? A study on the factuality evaluation capabilities of LLMs", *Proceedings of the Third Workshop on Natural Language Generation, Evaluation, and Metrics (GEM)*, Singapore, Association for Computational Linguistics, pp. 310-316, doi: 10.48550/arXiv.2311.00681.

Hajcak, G. and Foti, D. (2020), "Significance?. Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: an integrative review", *Psychophysiology*, Vol. 57 No. 7, e13570, doi: 10.1111/psyp.13570.

Hill, P.A. and Narine, L.K. (2023), "Ensuring responsible and transparent use of generative AI in extension", *Journal of Extension*, Vol. 61 No. 2, 13, doi: 10.34068/joe.61.02.13.

Lim, L.A., Dawson, S., Gašević, D., Joksimović, S., Fudge, A., Pardo, A. and Gentili, S. (2020), "Students' sense-making of personalised feedback based on learning analytics", *Australasian Journal of Educational Technology*, Vol. 36 No. 6, pp. 15-33, doi: 10.14742/ajet.6370.

Lin, C., Liu, X., Xv, G. and Li, H. (2021), "Mitigating sentiment bias for recommender systems", *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, doi: 10.1145/3404835.346294.

Markowitz, D.M. and Hancock, J.T. (2024), "Generative AI are more truth-biased than humans: a replication and extension of core truth-default theory principles", *Journal of Language and Social Psychology*, Vol. 43 No. 2, pp. 261-267, doi: 10.1177/0261927X231220404.

Samawi, F.S. and Al-Assaf, J.A.F. (2023), "Empowering early learners: the prospective impact of artificial intelligence on kindergarten education", *International Journal of Educational Sciences*, Vol. 42 Nos 1-3, pp. 54-62.

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M. and Olson, C.B. (2024), "Comparing the quality of human and ChatGPT feedback of students' writing", *Learning and Instruction*, Vol. 91, 101894, doi: 10.1016/j.learninstruc.2024.101894.

Wenling, L., Muhamad, M.M., Fakhruddin, F.M., Qiuyang, H. and Weili, Z. (2023), "Exploring the impact of emotional education in parent-child interactions on early childhood emotional intelligence development", *International Journal of Academic Research in Progressive Education and Development*, Vol. 12 No. 3, pp. 691-699, doi: 10.6007/IJARPED/v12-i3/18088.

Zhu, X., Zhang, Y., Wang, L., Feng, Q. and Li, W. (2024), "How does feedback type influence the cognitive engagement of students with varying levels of readiness for online learning? A network analytic approach", *2024 6th International Conference on Computer Science and Technologies in Education (CSTE)*, IEEE, pp. 309-314, doi: 10.1109/CSTE62025.2024.00064.

Corresponding author
Wei Wei can be contacted at: weiweitesting@hotmail.com