

**Generative AI or Human? Exploring the Role of AI- and Feedback Literacy in Student  
Feedback Perceptions**

Lena Fischer

Department of Psychology, University of Twente

Educational Psychology: 202200364

Bachelor's Thesis submitted in partial fulfilment of the requirements for the degree of

Bachelor of Science in Psychology

First Supervisor: Dr. Mohammadreza Farrokhnia

Second Supervisor: Dr. Chandan Dasgupta

Date: 25.06.2025

## **Generative AI or Human? Exploring the Role of AI- and Feedback Literacy in Student Feedback Perceptions**

### **Abstract**

As generative artificial intelligence (GenAI) becomes increasingly integrated into education, it is essential to comprehend students' perceptions of GenAI-generated feedback. This study examined how these perceptions are shaped by two key competencies: feedback literacy and AI literacy, focusing on university students' evaluations of feedback from both GenAI and human teachers.

In a study employing a within-subjects experimental design ( $N = 73$ ), participants assessed two pieces of feedback, one from a teacher and one from GenAI, both before and after the revelation of the feedback source. Results showed that evaluations were significantly affected by source awareness: GenAI-generated feedback was rated less favourably once its source was revealed, whereas teacher feedback was rated more favourably. Higher feedback literacy and AI literacy correlated with more positive evaluations of GenAI-generated feedback, with feedback literacy emerging as the more significant predictor.

While the study provides new perspectives on the influence of these literacies on feedback perceptions, its small, demographically limited sample and dependence on self-report data limit the ability to generalize findings. Future research should explore behavioural responses to GenAI-generated feedback in more diverse and authentic learning contexts.

*Keywords:* Feedback, Generative AI, Feedback Perception, Feedback Literacy, AI Literacy

## Table of Contents

<b>Introduction .....</b>	<b>5</b>
<i>Research Aim .....</i>	8
<b>Theoretical Framework .....</b>	<b>8</b>
<i>GenAI as a feedback source .....</i>	8
<i>Feedback perception .....</i>	9
<i>Feedback literacy .....</i>	11
<i>AI literacy .....</i>	12
<b>Summary and Research Questions .....</b>	<b>14</b>
<b>Methods .....</b>	<b>14</b>
<i>Design .....</i>	14
<i>Participants .....</i>	15
<i>Procedure .....</i>	16
<i>Materials .....</i>	17
Preparation .....	17
Feedback Literacy Questionnaire .....	17
AI Literacy Questionnaire .....	18
Essay and Feedback Pieces .....	18
Evaluation Questionnaire .....	19
<i>Data Analysis .....</i>	19
Preparation .....	19
Descriptives .....	20
RQ1: How do GenAI and human teachers, as feedback sources, influence students' perceptions of feedback? .....	20
RQ2: To what extent does students' feedback literacy impact their evaluation of GenAI-generated and teacher feedback? .....	21
RQ3: To what extent does students' AI literacy impact their evaluation of GenAI-generated feedback? .....	22
<b>Results .....</b>	<b>23</b>
<i>RQ1: How do GenAI and human teachers, as feedback sources, influence students' perceptions of feedback? .....</i>	24
<i>RQ2: To what extent does students' feedback literacy impact their evaluation of GenAI-generated and teacher feedback? .....</i>	28
<i>RQ3: To what extent does students' AI literacy impact their evaluation of GenAI-generated feedback? .....</i>	31
<b>Discussion .....</b>	<b>33</b>
<i>RQ1: How do GenAI and human teachers, as feedback sources, influence students' perceptions of feedback? .....</i>	33
<i>RQ2: To what extent does students' feedback literacy impact their evaluation of GenAI-generated and teacher feedback? .....</i>	36
<i>RQ3: To what extent does students' AI literacy impact their evaluation of GenAI-generated feedback? .....</i>	37
<i>Limitations and Recommendations for Future Research .....</i>	38
<i>Implications for Practice .....</i>	39

<i>Conclusion</i> .....	41
<b>References</b> .....	<b>43</b>
<b>Appendices</b> .....	<b>53</b>
AI statement .....	53
Consent Form for “Generative AI or Human? Exploring the Role of AI- and Feedback Literacy in Student Feedback Perceptions” .....	54
Feedback Literacy Scale.....	56
AI Literacy Scale .....	58
The Essay .....	61
Feedback 2: GenAI-generated Feedback .....	63
Feedback Prompt for ChatGPT .....	65
Feedback 1: Teacher Feedback .....	67
Feedback Evaluation Form .....	68
R Script.....	69
Assumption Check Plots .....	80

## Introduction

Feedback is crucial to the learning process because it highlights strengths and areas for improvement to achieve the best possible outcomes (Krackov, 2016). Traditionally, a student receives feedback from a human teacher, which has been shown to effectively enhance engagement (Gan et al., 2021), support self-regulation (Butler & Winne, 1995; Brown et al., 2016), and ultimately lead to better learning outcomes (Saihi et al., 2024). However, this process can be highly time-consuming for teachers, which might limit how frequently and how thoroughly they can provide feedback to students (Zacharias, 2007). As a result, exploring alternative feedback techniques, such as peer feedback or automated systems, might lessen teachers' workload while ensuring students receive timely and constructive feedback.

With recent advancements in technology, artificial intelligence (AI)-generated feedback has emerged as an alternative to traditional teacher-provided feedback, offering students immediate, automated, and personalized responses (Cavalcanti et al., 2021). A major advancement in AI-generated feedback is the use of Generative Artificial Intelligence (GenAI), a subfield of AI that makes use of machine learning algorithms to create human-like content, including text and images, based on existing datasets (Mittal et al., 2024). Given its accessibility and rapid development, GenAI is increasingly being integrated into education, functioning as a valuable tool to support both students and teachers (Saihi et al., 2024).

GenAI has the potential to enhance both learning and assessment processes in education by offering multifaceted tools that support both students and teachers. According to Chan & Hu (2023), it can serve as a useful study tool by generating ideas, synthesizing information, and summarizing complicated texts in a more accessible manner for students. Additionally, GenAI can provide instant, individualized feedback that fosters student engagement and helps improve academic work by pinpointing areas that require improvement (Zaman, 2023). Beyond learning support, GenAI can also optimize assessment processes by automating the grading of essays (Lee & Yoke, 2024; Wetzler et al., 2024) and delivering detailed, consistent, and scalable

feedback (Banihashem et al., 2024). Unlike traditional teacher-provided feedback, which can be time-consuming and subject to variability (Escalante et al., 2023), GenAI-generated feedback ensures greater consistency in evaluation (Lee & Yoke, 2024; Rad et al., 2023). By reducing the workload associated with grading, GenAI enables educators to allocate more time to lesson planning, administrative responsibilities, and creating a well-structured, student-centred learning environment (Shailendra et al., 2024; Gul et al., 2021).

However, despite these benefits, feedback from GenAI is not without drawbacks. It may sometimes lack pedagogical depth, emotional attunement, or contextual nuance; qualities that human teachers intuitively provide during the feedback process (Lee & Moore, 2024). Acknowledging both the benefits and corresponding limitations, it becomes increasingly important to reflect on how students assess and react to GenAI-generated feedback.

While GenAI has shown promise in providing feedback on students' assignments (Kohnke et al., 2025), the extent to which students engage with and apply this feedback is largely influenced by their perception of its content quality such as fairness, relevance, and factual accuracy, among other factors (Nazaretsky et al., 2024). If students do not perceive feedback as fair or reliable, they may be less likely to trust, engage with, or implement the suggested revisions (Henderson et al., 2019). On the other hand, scepticism about feedback, whether due to perceived biases, inaccurate information, or lack of personalization, can result in lower feedback uptake and limited impact on academic performance (Harrison et al., 2014). This issue becomes even more significant when GenAI is the source of feedback, as it raises persistent concerns about potential bias, inaccuracies, and ethical issues such as plagiarism and academic integrity (Chiu, 2023; Chan & Hu, 2023).

These concerns are closely related to how students interpret and evaluate GenAI-generated feedback, which is heavily influenced by their knowledge of and experience with the technology. In other words, students' past encounters with feedback, whether from teachers, peers, or automated systems, can influence how they interpret and respond to new sources of

feedback (Jensen et al., 2022). If students are familiar with receiving detailed, constructive, and tailored feedback from human instructors, they may be more critical of GenAI-generated feedback, perceiving it as less multifaceted or relevant (Ramu et al., 2023). In contrast, students who have had inconsistent or superficial feedback experiences may be more open to GenAI-generated feedback, especially if it provides structured, immediate, and actionable suggestions (Roe et al., 2024). Research by Nazaretsky et al. (2024) demonstrated that when students were unaware of the feedback source, they rated GenAI-generated feedback positively. However, once informed that the feedback was GenAI-generated, they preferred human feedback, suggesting an innate prejudice against GenAI. This is consistent with other research findings, which suggest that awareness of the feedback source can influence perceptions of its quality (Lim & Schmälzle, 2024; Raj et al., 2023).

Such a change in perspective can be attributed, in part, to students' levels of feedback literacy and AI literacy, which shape how they evaluate and engage with the received feedback (Jin et al., 2025; Richmond & Nicholls, 2024). Feedback literacy, defined as the ability to seek, interpret, and apply feedback effectively (Carless & Boud, 2018), plays a crucial role in students' ability to engage with feedback. Those with strong feedback literacy skills are more capable of deriving meaningful insights from feedback, regardless of its source, and incorporate it into their learning process (Carless & Winstone, 2020). Similarly, AI literacy encompasses the ability to understand, critically assess, and effectively interact with AI technologies (Ng et al., 2021). Students with higher AI literacy are more likely to recognize the strengths and limitations of GenAI-generated feedback, enabling them to make more informed judgments about its credibility and usefulness (Kelly et al., 2023).

## **Research Aim**

Despite the essential role that feedback literacy and AI literacy play in shaping students' perceptions of received feedback and, consequently, their feedback uptake, these literacies have largely been examined in isolation. Although feedback literacy and AI literacy have each been explored in relation to GenAI, no prior study has examined their respective influence on students' evaluations of both GenAI- and teacher-generated feedback within the same empirical context. Nazaretsky et al. (2024), for example, demonstrated that students' perceptions of the quality of feedback can shift once the source is revealed; yet they did not investigate the underlying competencies that shape these changes.

This study aims to address the abovementioned gap by exploring how students' feedback literacy and AI literacy influence their evaluations of feedback from two distinct sources: a human teacher and GenAI. By doing so, the findings will provide educators, instructional designers, and policymakers with insightful information and evidence-based recommendations to enhance students' engagement with feedback by fostering targeted literacies. Ultimately, a deeper understanding of these aspects could support a more effective integration of GenAI-generated feedback in educational settings, enabling students to critically evaluate and make meaningful use of the feedback they receive.

## **Theoretical Framework**

### **GenAI as a feedback source**

Recent studies highlight the growing role of GenAI in delivering educational feedback, emphasizing its immediacy, scalability, and personalization (Corbin et al., 2025). Compared to human feedback, GenAI is excelling in delivering prompt responses and decreasing instructional workload for teachers (Crompton, 2024). It has been widely adopted in areas such as automated essay grading, language learning, and skill development (Lee & Yoke, 2024).



These advancements make GenAI a practical choice for improving feedback in assessments that demand efficiency and scalability.

Despite these advantages, research suggests that GenAI-generated feedback may lack the relational and emotional depth that is associated with human feedback. While it can enhance student engagement and support self-regulated learning, it often does not establish trust or mutual recognition in the way human educators do (Lee & Moore, 2024). A recognition-based framework highlights that effective feedback is not just informational but relational (Corbin et al., 2025). Additionally, a student feedback engagement framework emphasizes that although GenAI fosters greater accessibility and reduces feedback anxiety, its effectiveness relies on students' ability to critically evaluate and apply the feedback received (Zhan et al., 2025). Ultimately, how students perceive feedback, whether from GenAI or human sources, significantly affects their willingness to engage with and act on it. The following section will explore how feedback perception shapes learning outcomes and how it can be optimized in the context of AI-assisted education.

### **Feedback perception**

Feedback perception refers to how individuals experience and respond to the feedback they receive, comprising their cognitive, emotional, motivational, and behavioural engagement (Van der Kleij & Lipnevich, 2020). This perception is shaped by various interconnected factors related to the feedback itself, including the quality, clarity, relevance, perceived sincerity, fairness, specificity, objectivity, and emotional tone (Strijbos et al., 2021; Woodrow et al., 2024; Fu & Nassaji, 2016).

Generally, detailed, specific, and constructive feedback is seen as more useful and promotes engagement, while vague or ambiguous comments are often dismissed or misinterpreted (Woodrow et al., 2024; Struyven et al., 2005). The credibility of the feedback

source also plays a major role (van de Ridder et al., 2014). Feedback from teachers or peers tends to carry relational value and trust (Heron et al., 2021), aspects that automated systems often struggle to replicate (Langer et al., 2022).

Additionally, the delivery style, including tone and phrasing, shapes emotional responses. A respectful, supportive tone can enhance motivation and receptiveness, even when providing critical feedback (Tanes et al., 2011). Timely feedback is also essential: prompt feedback enables learners to reflect and revise effectively, whereas delays can reduce its perceived usefulness (Rowe & Wood, 2009).

Further evidence suggests that the source of feedback directly influences how students engage with and act on it. Ruegg (2015) found that while students were more likely to attempt revisions based on teacher feedback, they often made more successful revisions based on peer feedback, possibly due to the increased clarity and relatability of peer comments. Teacher feedback, despite being detailed and specific, was more frequently misunderstood, indicating that perceived accessibility and student confidence are also important factors for understanding feedback.

Taken together, these observations show that feedback is not merely about the information that is communicated, but also about how it is received, interpreted, and acted upon by learners. Accordingly, fostering effective engagement with feedback requires more than improving its technical quality or delivery; it also demands equipping students with the capacity to make sense of and apply feedback constructively.

This need becomes increasingly critical as educational settings increasingly integrate GenAI technologies to enhance formative assessment. Although GenAI-generated feedback offers several advantages, such as immediacy, scalability, and consistency, learners may view it as less authentic, personalized, or trustworthy due to its automated nature (Nazaretsky et al., 2022). Furthermore, features such as the level of detail, tone, types of errors addressed, and the

situational context in which feedback is delivered all shape learners' interpretations (Fu & Nassaji, 2016).

As the use of GenAI in education continues to expand, understanding how students perceive and process this form of feedback, compared to feedback from teachers or peers, becomes increasingly important. Equally important, however, is recognizing that this change in feedback sources necessitates a corresponding development in learner competencies. Specifically, it underscores the need to build new forms of literacy, namely, *feedback literacy* and *AI literacy*, as explained next.

### **Feedback literacy**

Feedback literacy refers to the capabilities and behaviours that students need to engage with, make sense of, and effectively use feedback to support their learning (Dawson et al., 2023). It contains not only the skills required to respond to feedback but also the mindset needed to engage with it productively. Instead of being a passive recipient, a feedback-literate student actively seeks out feedback, critically evaluates its content, and applies it meaningfully to improve their work (Little et al., 2023). Students who demonstrate high levels of feedback literacy are more likely to benefit from the feedback process, as they possess the competence to interpret feedback accurately and translate it into concrete learning strategies (Carless & Boud, 2018). These learners tend to approach feedback as a constructive tool for development rather than as personal criticism. In contrast, students with low feedback literacy may struggle to derive meaningful benefit from feedback, often misinterpreting, dismissing, or misapplying the information provided (Carless & Winstone, 2020).

According to empirical research by Zhang, Schunn, and Wu (2024), students who possess higher levels of feedback literacy are better equipped to apply feedback they receive, especially when it comes to giving constructive peer feedback. However, the moderate strength

of this relationship suggests that feedback uptake is additionally shaped by motivational and emotional factors.

Next to peer feedback, feedback from teachers serves as a crucial source for developing feedback literacy. Wei, Sun, and Xu (2021) found that teacher feedback supports feedback literacy when it is not merely correcting mistakes but rather takes a more dialogic and actionable form. Senior students highly valued feedback that included concrete improvement suggestions, self-assessment opportunities, and alignment with personal goals.

Another source of feedback is AI tools. They offer immediate and detailed feedback, yet it might be difficult to interact with. Zhan and Yan (2025) found that students using ChatGPT often focused on extracting useful information and evaluated GenAI-generated feedback by comparing it to specific reference points, such as their own writing, previous teacher comments, and IELTS scoring criteria (official guidelines for assessing writing based on task response, coherence, vocabulary, and grammar). However, they showed weaker metacognitive regulation, which sometimes led to superficial engagement. Although GenAI-generated feedback often helps students feel less stressed about their academic work, students must critically assess its trustworthiness (Tubino & Adachi, 2022). Developing skills like evaluative judgment, emotional reflexivity, prompt engineering, and ethical decision-making is crucial to prevent overreliance and uphold academic integrity (Zhan & Yan, 2025), highlighting the need for *AI literacy* besides traditional feedback literacy.

## **AI literacy**

AI literacy can be defined as a set of skills and knowledge that enables individuals to critically evaluate AI technologies, interact and collaborate effectively with AI systems. These skills also enable people to apply AI tools across various contexts, including everyday life, work, and online environments (Wang et al., 2024). AI literacy emphasizes an understanding

of how AI functions, its societal and personal impacts, and how to engage with it in an informed and responsible manner, without requiring technical expertise in AI development (Long & Magerko, 2020). Individuals with a technical background or prior experience with AI often demonstrate higher levels of AI literacy (Hornberger et al., 2023). They are typically more capable of recognizing AI-generated content and are more inclined to question its authenticity and accuracy, rather than accepting it without scepticism simply because it appears well-produced (Kelly et al., 2023). Furthermore, those with strong AI literacy are better capable of using AI tools strategically, which enhances their learning, creativity, and productivity. They are also more likely to have a strong understanding of ethical considerations, guiding them to prevent practices such as plagiarism and uphold standards of academic integrity (Hutson, 2024).

Importantly, AI literacy plays a significant role in shaping students' perceptions of grades, feedback, and other evaluative input provided by GenAI systems. For instance, students with higher AI literacy are more likely to approach GenAI-generated feedback critically, examining its credibility, specificity, and relevance before acting upon it (Yan et al., 2024). They are also better at identifying limitations, such as generic phrasing or factual inaccuracies, and supplementing AI's suggestions with their judgment. In contrast, students with lower AI literacy may either over-rely on GenAI outputs without sufficient critical reflection or dismiss them entirely out of mistrust or unfamiliarity; both are tendencies that are observed in learners with limited exposure to AI tools (Tzirides et al., 2024).

This ability to critically evaluate is particularly crucial when feedback influences academic achievement or skill development. Lacking sufficient AI literacy, students may struggle to engage effectively with GenAI-generated feedback, potentially misinterpreting its purpose or failing to fully utilize it. Consequently, AI literacy acts as both a cognitive and ethical filter, shaping how students perceive, interpret, and apply feedback generated by AI tools. Additionally, it promotes responsible use, motivating students to respect academic integrity by acknowledging and resolving AI systems' underlying limitations.

## Summary and Research Questions

The increasing presence of GenAI in education raises important questions about how students perceive and use GenAI-generated feedback. This perception is shaped not only by the feedback's source but also by students' competencies, particularly feedback literacy and AI literacy. Although previous research has addressed these literacies separately, no study to date has simultaneously examined their influence on how students evaluate feedback from both GenAI and human teachers in a direct comparison.

Given this foundation, the current study aims to explore how students interpret feedback from GenAI and teachers, and how their literacy levels inform this process.

Thereby, this study aims to provide answers to the following questions:

**RQ1.** How do GenAI and human teachers, as feedback sources, influence students' perceptions of feedback?

**RQ2.** To what extent does students' feedback literacy impact their evaluation of GenAI-generated and teacher feedback?

**RQ3.** To what extent does students' AI literacy impact their evaluation of GenAI-generated feedback?

## Methods

### Design

The current study employed a quantitative, cross-sectional, correlational research design to examine how students' feedback literacy and AI literacy influence their perceptions and evaluations of feedback from GenAI and teachers. Data were collected through an online survey comprising standardized instruments measuring feedback literacy, AI literacy, and feedback

perception. Participants ( $N = 73$ ) were recruited via convenience and snowball sampling, including outreach through the University of Twente's internal SONA recruitment platform. Apart from that, the study was uploaded to the SurveySwap website, where users can promote their surveys and support each other by completing surveys in exchange for credits. This creates a mutual system of help.

## Participants

A total of 73 participants were recruited, with ages ranging from 18 to 40 years ( $M = 22.6$ ,  $SD = 3.10$ ). Most identified as female and reported German or Dutch nationality. A full demographic summary is presented in Table 1. All participants provided informed consent, and both confidentiality and anonymity were maintained throughout the whole research process. The study was approved by the Ethics Committee of the University of Twente.

**Table 1**

*Demographic Summary of Participants in a Table*

Variable	Category	N	%
<b>Age (in years)</b>	Mean (SD)	22.6	-
	Range	18-40	-
<b>Gender</b>	Female	57	78.08
	Male	15	20.55
	Non-binary/third gender	1	1.37
<b>Nationality</b>	German	34	46.58
	Dutch	19	26.03
	Other	20	27.39

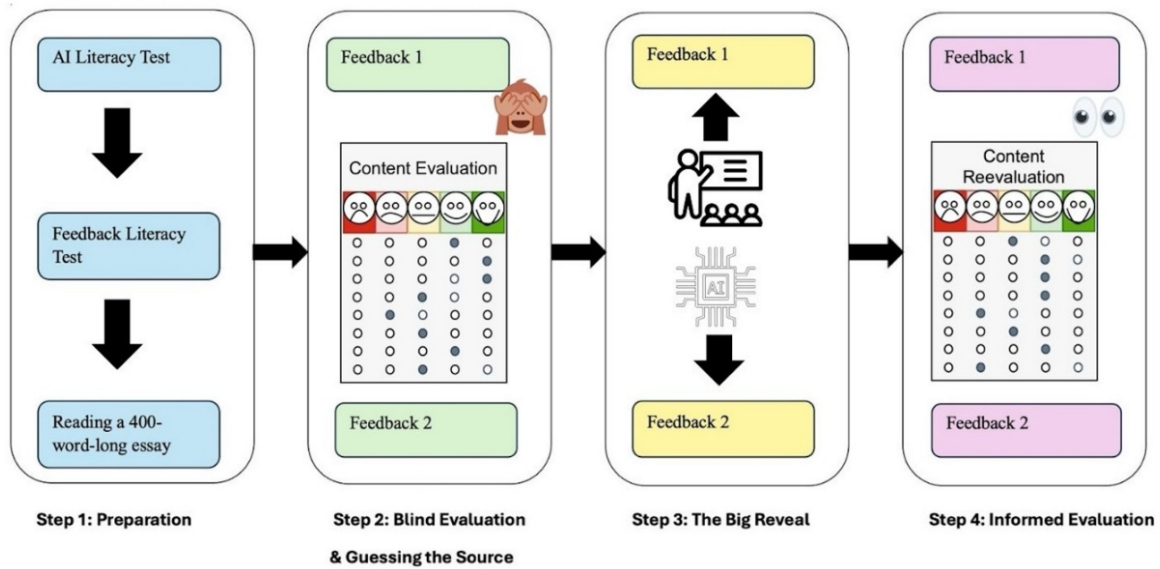
## **Procedure**

The study was conducted entirely online in a single session lasting approximately/at most 30 minutes (Figure 1). Upon accessing the survey link, participants were immediately directed to the informed consent form. After providing consent, they answered three sociodemographic questions related to gender, age, and nationality.

Subsequently, participants completed standardized assessments measuring their feedback literacy and AI literacy. They were then presented with a 400-word argumentative essay accompanied by a piece of feedback (teacher-generated). After reviewing the feedback, participants evaluated it using a structured questionnaire. This process was repeated for a second piece of feedback (Gen-AI generated) on the same essay. At this stage, the sources of the feedback were not disclosed.

After the initial evaluations, participants were asked to identify the presumed source of each feedback item by selecting from predefined options: peer, teacher, or GenAI. Once the actual sources were revealed, they were asked to reassess both feedback pieces using the same questionnaire as before. The survey concluded with a thank-you message for their participation.



**Figure 1***The Procedure of the Study***Materials***Preparation*

Participants used a digital device of their choice (mobile phone, laptop) to fill out the survey that was created and answered in Qualtrics. Before proceeding to the main survey, participants were presented with an informed consent form in Qualtrics. This form consists of the study procedure and participants' rights, including confidentiality, the right to withdraw at any time, and the opportunity to ask the researcher questions (Appendix B).

*Feedback Literacy Questionnaire*

The first questionnaire, which intends to measure feedback literacy derived from Dawson et al. (2024). It includes 24 items that are answered on a 6-point Likert scale. It contains 5 subscales, namely seek feedback information, make sense of information, use feedback

information, provide feedback information, and manage affect. The questionnaire demonstrates acceptable reliability and a strong initial validity. Appendix C contains the full instrument.

### *AI Literacy Questionnaire*

The second questionnaire testing the AI literacy of the participants was adapted from the article of Wang et al. (2024). It contains 28 statements that are answered on a 7-point Likert scale. The items measure the awareness, usage, evaluation, and ethics of AI. Additionally, they measure subjective norms, perceived behavioural control, attitudes, and behavioural intention to use AI. Demonstrating good internal consistency, convergent validity, and discriminant validity, the questionnaire is appropriate for measuring AI literacy in university student populations. The complete version is presented in Appendix D.

### *Essay and Feedback Pieces*

The essay (Appendix E) was written by the researcher, and the GenAI-generated feedback (Appendix F) was generated using ChatGPT-4o with an Auto-Chain of Thought prompt. This prompt was designed to guide the model through the task by encouraging step-by-step reasoning, similar to how a human would provide feedback. ChatGPT was instructed to take on the role of a university-level writing instructor, evaluating a 400-word argumentative essay based on specific criteria (structure, clarity of position, counterarguments, and conclusion). The prompt also included example input and output to illustrate what high-quality feedback might look like (Appendix G). The teacher's feedback (Appendix H) was written by a teacher.

### *Evaluation Questionnaire*

The questionnaire designed to evaluate the feedback content was produced by Nazaretsky et al. (2024) and measured eight criteria, namely preciseness, fairness, factuality, relevance, information, application, authenticity, and sincerity, on a 5-point Likert scale (Appendix I). These criteria are grounded in prior research and grouped into three broader dimensions: *Objectivity*, *Usefulness*, and *Genuineness*. *Objectivity* encompasses preciseness, fairness, and factuality, ensuring that feedback accurately reflects student performance. *Usefulness*, which includes relevance, informativeness, and applicability, is particularly crucial in shaping effective feedback engagement, as emphasized by Henderson et al. (2019) and Shute (2008). These researchers highlight that the degree to which feedback is actionable and informative strongly influences students' ability to benefit from it.

*Genuineness*, represented by authenticity and sincerity, ensures that feedback is delivered in a constructive and supportive manner rather than as judgment. This dimension, along with *Objectivity*, is further supported by Evans (2013), who argues that both factors are central to feedback quality as they foster students' ability to internalize and act upon the feedback they receive. Moreover, the importance of *Objectivity* and *Usefulness* is strongly reinforced by Hatti and Timperley's (2007) model, which underscores how accurate, goal-oriented feedback contributes to students' understanding of their progress and next steps.

Given these theoretical foundations, the questionnaire was chosen to ensure a structured and comprehensive assessment of students' feedback perceptions.

## **Data Analysis**

### *Preparation*

For conducting the data analysis, the data sheet was first transferred from Qualtrics into Excel. Survey preview and invalid responses were excluded, as were responses from

participants who did not complete the entire questionnaire. Subsequently, all qualitative responses were converted into their corresponding numerical values. The feedback literacy scale was coded as follows: 1 = Never, 2 = Almost never, 3 = Rarely, 4 = Sometimes, 5 = Almost always, 6 = Always. For the AI literacy scale, responses were coded: 1 = Strongly Disagree, 2 = Disagree, 3 = Somewhat Disagree, 4 = Neutral, 5 = Somewhat Agree, 6 = Agree, 7 = Strongly Agree. Feedback evaluation responses were transformed using the following scale: 1 = Strongly Disagree, 2 = Somewhat Disagree, 3 = Neither Agree nor Disagree, 4 = Somewhat Agree, 5 = Strongly Agree. Next, a mean score per participant was calculated across each respective section.

### *Descriptives*

In the following step, the data was imported into R Studio (R) for statistical analysis. Descriptive analyses were conducted on the demographic variables, including gender, age, and nationality. Measures of central tendency were calculated for age, specifically the range, mean, and standard deviation. Additionally, frequency distributions and percentages were computed for gender and nationality to provide a summary of the sample's composition. This information can be found in the Participants section. The mean and standard deviation of the feedback and AI literacy scores were computed, as well as the evaluation scores of both feedback pieces before and after the reveal of the sources. Appendix J contains the entire R script.

### *RQ1: How do GenAI and human teachers, as feedback sources, influence students' perceptions of feedback?*

For RQ1, to examine the impact of feedback source and source awareness on students' perceptions of feedback quality, a  $2 \times 2$  within-subjects repeated measures ANOVA was conducted, after the assumptions of normality, linearity, and homoscedasticity were evaluated

and considered reasonably met. The within-subject factors were Feedback Source (GenAI vs. Teacher) and Awareness (Blind vs. Known). Each participant rated feedback in four conditions: (1) GenAI-Blind, (2) GenAI-Known, (3) Teacher-Blind, and (4) Teacher-Known. The data were structured in RStudio so that each participant had four separate variables corresponding to these conditions. A bar plot was created to visualize the perceived quality of feedback by source and awareness, summarizing the mean ratings across the conditions.

Additionally, for both feedback texts, participants were asked to identify the perceived source by selecting one of three options: GenAI, teacher, or peer. Another bar chart was produced to display the number of correct and incorrect source identifications for each feedback type.

To investigate the relationship between participants' ability to correctly identify the feedback source and their perceptions of feedback quality, exploratory paired-samples t-tests were conducted on three subgroups: (1) participants who correctly identified both feedback sources, (2) participants who correctly identified the teacher feedback, regardless of their GenAI-generated feedback guess, and (3) participants who correctly identified the GenAI-generated feedback, regardless of their teacher feedback guess. For each subgroup, mean feedback ratings were compared, and a grouped bar plot with standard error bars was generated to visualize differences in perceived quality.

*RQ2: To what extent does students' feedback literacy impact their evaluation of GenAI-generated and teacher feedback?*

To explore how students' feedback literacy influenced their evaluation of both feedbacks, several analyses were conducted, after the assumptions of linearity, homoscedasticity, and normality of residuals for both regression models (GenAI and teacher) were evaluated and found to be reasonably met. First, Pearson correlation coefficients were

computed between standardized feedback literacy scores and feedback ratings under both the Blind and Known source conditions. The feedback literacy scores were standardized to be able to directly compare to AI literacy scores, as both variables were measured on different Likert scales. Standardization ensured that their relative predictive strength could be interpreted accurately.

In the Known condition, simple linear regression analyses were run to test whether feedback literacy significantly predicted feedback evaluations for both feedback sources. This approach enabled an examination of how differences in feedback literacy influence students' perceptions of different feedback sources within real-world educational contexts, which enhances the ecological validity and interpretability of the results.

Additionally, to assess group-level effects, participants were split into high and low feedback literacy groups via a median split. Their feedback ratings were then compared using Welch Two Sample t-tests. This test was selected as it does not assume equal variances between groups, making it a more reliable choice when sample sizes are unbalanced or when variances between groups might differ. It ensured that the statistical comparisons were robust and accounted for potential discrepancies in variance between high- and low-literacy student groups. Finally, scatter plots with fitted regression lines were generated to visually illustrate the relationships between feedback literacy and feedback evaluations.

*RQ3: To what extent does students' AI literacy impact their evaluation of GenAI-generated feedback?*

For RQ3, Pearson correlations were computed between standardized AI literacy scores and GenAI-generated feedback ratings across both Blind and Known conditions. As in RQ2, AI literacy scores were standardized to enable direct comparability with feedback literacy, ensuring that predictive strength could be interpreted precisely. In the Known condition, for

the same reason as before, a simple linear regression analysis was conducted to determine the predictive strength of AI literacy on feedback evaluations. Prior to analysis, assumption checks were conducted to verify linearity, homoscedasticity, and normality of residuals for the regression model.

Afterward, participants were categorized into high and low AI literacy groups based on a median split. To compare feedback ratings between these groups, the Welch Two Sample t-test was used since it does not require the assumption of equal variances across groups, making it better suited for situations with unequal sample sizes or when the variability between groups is inconsistent. This ensured the robustness of statistical comparisons, accounting for possible disparities in variance between high- and low-literacy student groups. Lastly, a scatterplot visualized the relationship between AI literacy and GenAI-generated feedback evaluations in the Known source condition.

## **Results**

Table 2 shows an overview of the mean scores and standard deviations of feedback literacy, AI literacy, and evaluation of GenAI-generated and teacher feedback before and after the source was revealed.

**Table 2**

*Descriptive Data of Feedback literacy, AI literacy, and Evaluation of Gen-AI Generated and Teacher Feedback Before and After the Source Reveal*

<b>Variable</b>	<b>Mean (M)</b>	<b>Standard Deviation (SD)</b>
Feedback Literacy Score	4.80	0.57
AI Literacy Score	5.18	0.71
Evaluation GenAI-generated Feedback (before the reveal)	4.05	0.69
Evaluation GenAI-generated Feedback (after the reveal)	3.84	0.81
Evaluation Teacher feedback (before the reveal)	4.04	0.67
Evaluation Teacher feedback (after the reveal)	4.20	0.68

*Note.* Feedback Literacy is the average participant score across all Feedback Literacy items and was measured on a 6-point Likert scale. AI Literacy is the average participant score across all AI Literacy items and was measured on a 7-point Likert scale. Evaluation of the feedback is the average participant score across the dimension's fairness, relevance, factuality, preciseness, information, applicability, authenticity, and sincerity, and was measured on a 5-point Likert scale.

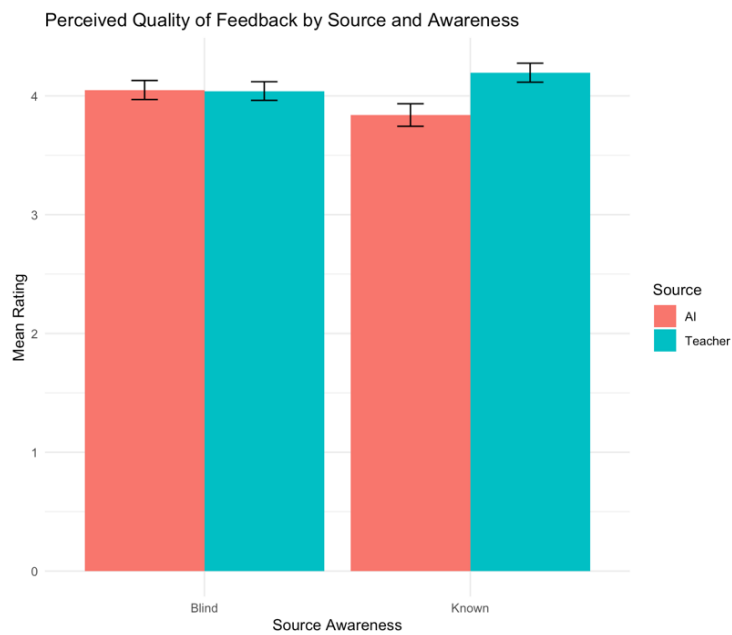
### **RQ1: How do GenAI and human teachers, as feedback sources, influence students' perceptions of feedback?**

A  $2 \times 2$  within-subjects repeated measures ANOVA was performed to analyse the effects of feedback source (GenAI vs. Teacher) and source awareness (Blind vs. Known) on students' evaluation of feedback quality. Before analysis, assumption checks confirmed that linearity and homoscedasticity were reasonably met based on residual plots (Appendix K, Figure K.1). Even though the Shapiro-Wilk test indicated a significant deviation from normality ( $W = 0.96, p > .001$ ), Q-Q plot inspection showed that residuals were approximately normally distributed, with only slight deviations at the extremes. Given the balanced, within-subjects



design and the ANOVA's robustness to moderate non-normality, parametric assumptions were considered reasonably met, and the analysis was carried out.

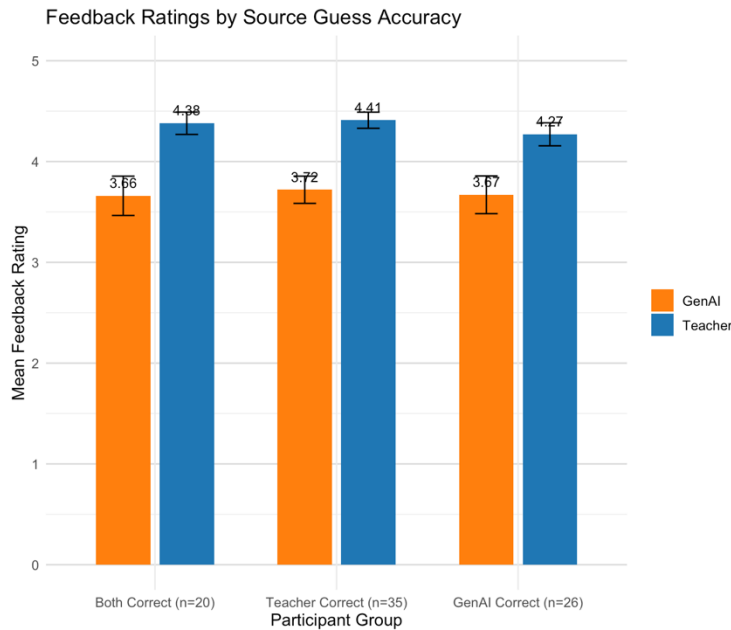
The analysis revealed no significant main effect of feedback source ( $F(1, 72) = 3.14, p = .08$ ), suggesting that, overall, students did not rate GenAI-generated and teacher feedback differently when evaluating it without knowing its source. Similarly, there was no significant main effect of source awareness ( $F(1, 72) = 0.82, p = .37$ ), indicating that knowing the source alone did not consistently affect feedback ratings. However, there was a significant interaction between feedback source and awareness ( $F(1, 72) = 20.64, p < .001$ ), with a modest effect size (generalized eta squared = 0.016), showing that the effect of feedback source on evaluation scores depended on whether the source was known. As shown in the accompanying bar plot (Figure 2), mean ratings varied across conditions: when the source was unknown (Blind), GenAI-generated and teacher feedback were rated similarly, but once the source was revealed (Known), ratings for feedback from GenAI dropped while ratings for teacher feedback increased. This crossover pattern illustrates that students' evaluations were impacted by the feedback content as well as by their perceptions of the source, with GenAI-generated feedback scoring lower when its source was disclosed.

**Figure 2***Perceived Quality of Feedback by Source and Awareness*

Moreover, as shown in Figure 3, 35 participants correctly identified the teacher-generated feedback (ca. 48%), while 38 did not. For the GenAI-generated feedback, 26 participants identified the source correctly (ca. 36%), whereas 47 responded incorrectly.

**Figure 3***Participant Identification of Feedback Source*

To examine whether the observed preference for teacher feedback over GenAI-generated feedback was influenced by participants' ability to identify the source correctly, separate analyses for participants who correctly guessed the source of feedback before disclosure were conducted. When restricting the sample to participants who correctly guessed both sources ( $n = 20$ ), a paired-samples t-test showed that teacher feedback ( $M = 4.38$ ) was rated significantly higher than GenAI-generated feedback ( $M = 3.66$ ),  $t(19) = 3.49$ ,  $p = .002$ . Similarly, among participants who correctly identified the teacher feedback regardless of their GenAI-generated feedback guess ( $n = 35$ ), teacher feedback ( $M = 4.41$ ) continued to be significantly preferred over GenAI-generated feedback ( $M = 3.72$ ),  $t(34) = 4.97$ ,  $p < .001$ . Also, participants who correctly identified the GenAI-generated feedback, regardless of their teacher feedback guess ( $n = 26$ ), rated teacher feedback ( $M = 4.27$ ) significantly higher than GenAI-generated feedback ( $M = 3.67$ ),  $t(25) = 3.00$ ,  $p = .006$ . These findings are visually summarized in Figure 4.

**Figure 4***Perceived Feedback Quality Based on Source Identification Accuracy*

**RQ2: To what extent does students' feedback literacy impact their evaluation of GenAI-generated and teacher feedback?**

To explore the relationship between students' feedback literacy and their evaluation of feedback, a series of correlation and regression analyses were conducted. Before analysis, assumptions of linearity, normality of residuals, and homoscedasticity were evaluated and found to be reasonably met (Appendix K, Figures K.2-K.5). The Shapiro-Wilk test indicated that the residuals for the GenAI model were normally distributed ( $W = 0.9769, p = .200$ ), while the test for the teacher model revealed a significant deviation from normality ( $W = 0.8578, p < .001$ ). However, Q-Q plot inspections showed that residuals were approximately normally distributed in both models, with only slight deviations at the extremes. Given these results and the robustness of linear regression to moderate non-normality, both models were deemed appropriate for analysis.

Pearson correlation analyses demonstrated a significant positive association between feedback literacy and students' evaluations of feedback in both GenAI-generated and teacher conditions. Specifically, evaluations of feedback from GenAI revealed significant correlations in both the Blind ( $r = .37, p = .001$ ) and Known ( $r = .36, p = .002$ ) conditions, indicating that higher levels of feedback literacy were consistently associated with more favourable evaluations of GenAI-generated feedback, regardless of source awareness. Similarly, evaluations of teacher feedback showed significant correlations in the Blind ( $r = .30, p = .010$ ) and Known ( $r = .33, p = .005$ ) conditions, suggesting a comparable pattern in which greater feedback literacy was associated with more positive perceptions of teacher feedback.

To further assess this relationship, simple linear regression analyses were conducted using feedback literacy as a predictor of GenAI-generated feedback evaluation in the Known condition. The model was statistically significant ( $F(1, 71) = 10.26, p = .002$ ), suggesting that feedback literacy meaningfully predicted students' perceptions of GenAI-generated feedback. Specifically, higher feedback literacy scores were associated with higher evaluation ratings ( $\beta = 0.51, p = .002$ ). The standardized beta coefficient indicated a moderate effect size ( $\beta = 0.51, p = .002$ ), and the model accounted for approximately 13% of the variance in feedback evaluations ( $R^2 = 0.13$ ). Similarly, the regression model predicting teacher feedback evaluations was significant ( $F(1, 71) = 8.53, p = .005$ ), with feedback literacy positively influencing students' perceptions of teacher feedback ( $\beta = 0.45, p = .005$ ). This model explained approximately 11% of the variance ( $R^2 = 0.11$ ), suggesting a similar but slightly weaker effect compared to GenAI-generated feedback.

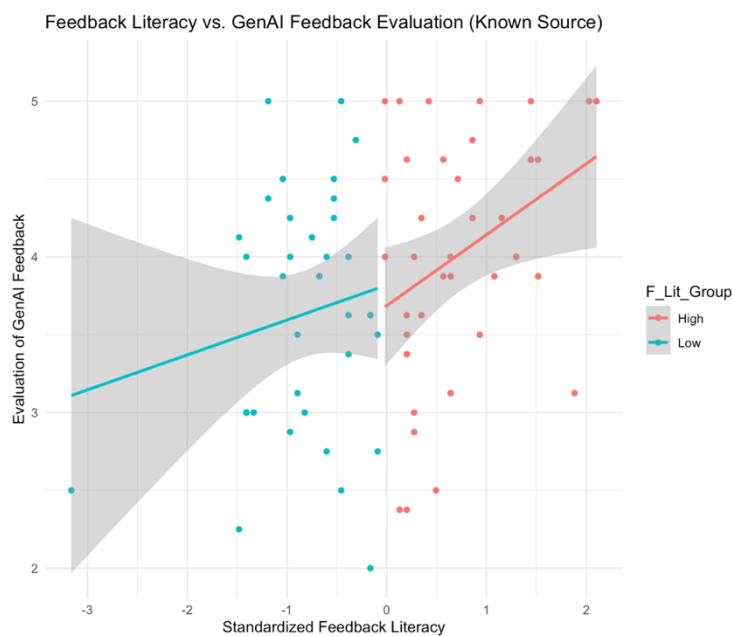
To compare feedback literacy effects across evaluation conditions, a series of Welch Two Sample  $t$ -tests examined differences in feedback ratings between students with high versus low feedback literacy. In the GenAI Known condition, students classified as having high feedback literacy reported significantly higher evaluations ( $M = 4.02$ ) compared to those with lower literacy ( $M = 3.63$ ),  $t(69.60) = 2.12, p = .038$ . A similar effect emerged in the teacher

feedback condition, with high-literacy students assigning significantly higher ratings ( $M = 4.35$ ) than their lower-literacy peers ( $M = 4.01$ ),  $t(66.95) = 2.22$ ,  $p = .030$ .

Scatter plots with fitted regression lines (Figures 5 and 6) visually illustrate the positive trend between feedback literacy and feedback evaluations, confirming that students with greater literacy tend to rate feedback more favourably across both sources.

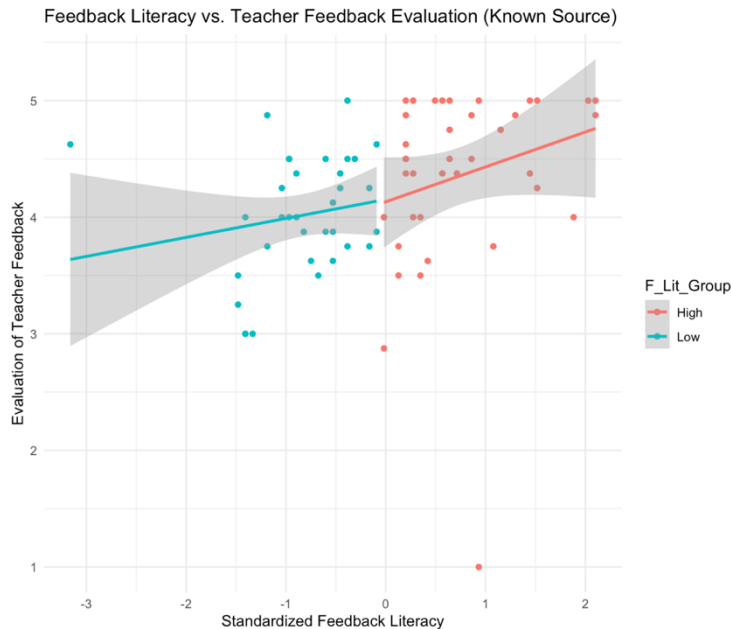
**Figure 5**

*Scatterplot Representing the Relationship between Feedback Literacy and GenAI-generated Feedback Evaluations in the Known Source Condition*



**Figure 6**

*Scatterplot Representing the Relationship between Feedback Literacy and Teacher Feedback Evaluations in the Known Source Condition*



### **RQ3: To what extent does students' AI literacy impact their evaluation of GenAI-generated feedback?**

In response to RQ3, the relationship between AI literacy and students' evaluations of GenAI-generated feedback was examined. Before conducting the analysis, residual diagnostics were reviewed to assess assumptions of linearity, homoscedasticity, and normality. Plots indicated that the model adhered to the former two assumptions (Appendix K, Figures K.6-K.7). Although the Shapiro-Wilk test suggested a deviation from normality ( $W = 0.963, p = .033$ ), the Q-Q plot showed only slight deviations at the distribution's ends, and residuals were otherwise well-aligned with expectations. Given the model's resilience to modest normality violations, the regression analysis was carried out.

Pearson correlation analyses revealed a significant positive association in both the Blind ( $r = .39, p = .0006$ ) and Known ( $r = .37, p = .0014$ ) conditions, indicating that higher AI literacy

was consistently linked to more favourable evaluations of GenAI-generated feedback, regardless of source awareness.

A simple linear regression was conducted to determine whether AI literacy significantly predicted students' evaluations of GenAI-generated feedback in the Known condition. The model was statistically significant ( $F(1, 71) = 11.02, p = .0014$ ), demonstrating that higher AI literacy scores were associated with higher feedback ratings. The standardized beta coefficient indicated a moderate effect size ( $\beta = 0.42, p = .0014$ ), and the model accounted for approximately 13.4% of the variance ( $R^2 = 0.134$ ), suggesting a relatively small but meaningful effect.

To further examine the impact of AI literacy, participants were categorized into high- and low-AI literacy groups based on a median split. A Welch Two Sample t-test indicated a significant difference in feedback evaluations between the groups ( $t(69.13) = 2.82, p = .006$ ). Students in the high AI literacy group rated GenAI-generated feedback significantly higher ( $M = 4.09$ ) compared to those in the low AI literacy group ( $M = 3.58$ ).

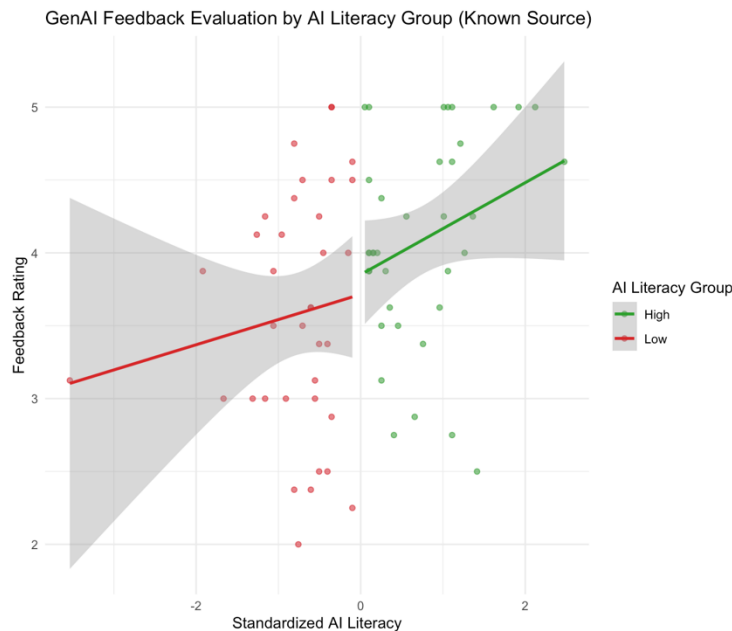
An independent samples *t*-test revealed a significant difference in feedback evaluations between the two groups ( $t = 2.82, p = .006$ ). Students in the high-AI literacy group rated GenAI-generated feedback more favourably ( $M = 4.09$ ) compared to those in the low-literacy group ( $M = 3.58$ ).

A scatter plot with a fitted regression line (Figure 7) visually confirms these findings, illustrating a clear positive trend between AI literacy and GenAI-generated feedback evaluations in the Known condition.



**Figure 7**

*Scatterplot Representing the Relationship Between AI Literacy and GenAI-generated Feedback Evaluations in the Known Source Condition*



## Discussion

The purpose of the current study was to investigate how students evaluate feedback provided by GenAI and teachers, and to what extent their levels of feedback literacy and AI literacy shape these evaluations. Below, each finding will be discussed in relation to the relevant research question.

### **RQ1: How do GenAI and human teachers, as feedback sources, influence students' perceptions of feedback?**

In response to RQ1, the findings show that students' perceptions of feedback were initially similar regardless of whether it was generated by GenAI or provided by a human teacher. This implies that, without source information, the two feedback pieces were primarily evaluated based on their content, rather than on their origin. However, once the source was

disclosed, a shift occurred in students' evaluations; teacher feedback was rated more positively while GenAI-generated feedback received lower ratings, aligning with the findings of Nazaretsky et al. (2024). They argue that such patterns among students reflect a strong preference for human teachers as feedback providers.

Importantly, this preference for teacher feedback persisted even among participants who had accurately identified the source of the feedback prior to disclosure. Whether participants recognized both sources correctly or only one, they still rated teacher feedback more favourably once the actual sources were confirmed. This subgroup analysis was essential in showing that source misattribution alone does not account for the effect. The findings indicate that it is not merely guessing incorrectly that skews evaluations, but rather source disclosure itself activates or reinforces underlying biases about credibility and trustworthiness. In other words, correct identification does not protect against bias; simply knowing the source appears to be sufficient to influence judgment.

This observation leads to the question of what drives the more positive perception of teacher feedback after the source becomes known. For one, source disclosure probably triggers pre-existing beliefs and biases, leading students to attribute greater expertise, care, and diligence to teachers while perceiving GenAI as less personalized or credible (Lee & Song, 2024). This reflects a type of authority bias, in which human feedback is given greater trust due to the social and institutional role of teachers (Raviv et al., 2003). Moreover, students may view teacher feedback as more personalized and thoughtful, whereas GenAI-generated feedback might be seen as generic or automated, even if the content is the same (Winstone et al., 2019).

Additionally, evaluations can be influenced by heuristic thinking: if the source is unknown, students may focus on the content, but when the source is revealed, rapid judgments derived from mental shortcuts, like associating AI with a diminished reliability, can bias their evaluation (Dale, 2015). Ultimately, a lack of familiarity with AI tools or uncertainty about how GenAI works may play a part in generating scepticism or reduced trust regarding feedback

produced by GenAI (Chan & Hu, 2023). Together, these patterns highlight that feedback evaluations were shaped not only by the content but also by the perceptions of source credibility, though the interaction effect was modest in magnitude.

Beyond these evaluation patterns (favouring teacher over GenAI-generated feedback), students were also more accurate at identifying the source of teacher feedback compared to that of GenAI-generated feedback. While teacher feedback was recognized with greater accuracy (48%), students struggled to attribute GenAI-generated feedback (36%) accurately. Nearly one-third of participants incorrectly guessed that it came from a teacher, about one-third credited it to a peer, and only about one-third correctly identified it as GenAI.

Several factors may explain this identification bias. First, students are naturally familiar with teacher communication styles, including their phrasing and structure, making human feedback more recognizable (Rawal, 2023; Grønli et al., 2025). Secondly, teacher feedback frequently includes contextually relevant references, instructional nuance, or emotional tone that feels more personal and relational (Huang et al., 2025). In contrast, GenAI-generated feedback may lack these relational indicators, despite being accurate and well-structured (Lee & Moore, 2024; Yan et al., 2024).

Another possible explanation could be incorrect expectations about how AI sounds. Students might think that GenAI is more robotic or flawed than it is, which can result in misconceptions when they attempt to recognize its output (Chan & Hu, 2023). Furthermore, the novelty and ambiguity of GenAI may lead to confusion. Students might lack a clear mental model of GenAI-generated feedback, making it harder for them to recognize it accurately (Vora et al., 2023).

In sum, the research question can be answered as follows: Students' perceptions of feedback are influenced by source awareness, leading them to rate teacher feedback more favourably than GenAI-generated feedback, even when they correctly identified the source

beforehand. However, the effect was relatively small and should therefore be interpreted with caution in terms of its practical implications.

**RQ2: To what extent does students' feedback literacy impact their evaluation of GenAI-generated and teacher feedback?**

Answering RQ2, the results indicate that feedback literacy meaningfully predicts students' perceptions of both feedback sources. Specifically, students possessing higher feedback literacy consistently provided more favourable evaluations of GenAI-generated feedback, particularly when the source was disclosed, compared to students with lower feedback literacy. Likewise, higher feedback literacy was associated with more positive evaluations of teacher feedback, regardless of whether the source was known or unknown, though this effect was slightly weaker than for GenAI-generated feedback. This pattern can be attributed to the fact that students who possess higher feedback literacy are more capable of critically examining feedback based on its informational quality rather than being influenced by its source (Malecka et al., 2020). Their enhanced skills enable them to interpret and utilize feedback effectively, allowing them to see the worth in comments from both teachers and GenAI, even when others may regard feedback generated by AI as less credible.

Feedback literacy thus also helps reduce source bias; while students with lower feedback literacy may rely on stereotypes or preconceived notions (such as viewing AI feedback as less trustworthy), more literate students assess feedback based on its content quality, resulting in more balanced and positive evaluations across different sources (Di Zhang et al., 2023). Moreover, the slightly stronger positive effect of feedback literacy on evaluations of GenAI-generated feedback suggests that these students are more flexible and open to engaging with novel feedback forms (Han & Xu, 2019). They can manage the uncertainty surrounding GenAI-

generated feedback, applying the same critical and constructive approach they use for human feedback. This fosters greater trust and acceptance of GenAI input.

Overall, students' feedback literacy significantly shapes their evaluation of both teacher and GenAI-generated feedback, promoting more positive and balanced perceptions by encouraging focus on feedback quality rather than its origin.

### **RQ3: To what extent does students' AI literacy impact their evaluation of GenAI-generated feedback?**

Regarding RQ3, the results show that students with greater AI literacy consistently evaluate GenAI-generated feedback more positively, regardless of whether the source is disclosed. This suggests that familiarity with AI technologies enhances students' understanding of the strengths and limitations of automated feedback, reducing scepticism and promoting trust in its credibility and usefulness (Kelly et al., 2023). Higher AI literacy likely enables students to critically engage with AI-generated content, focusing more on its informational value than on preconceived biases about non-human sources (Casal-Otero et al., 2023). However, it is worth mentioning that while AI literacy has a consistent positive effect on evaluations of GenAI-generated feedback, feedback literacy seems to have a more significant overall impact, particularly in determining how students evaluate feedback regardless of its source. This differentiation highlights the complementary yet distinct functions of AI literacy (focused on understanding AI systems) and feedback literacy (focused on assessing feedback quality) in shaping students' interactions with both teacher and GenAI-generated feedback.

Therefore, students' AI literacy plays a significant role in building understanding and trust regarding GenAI-generated feedback, which can lead to more positive evaluations of it. However, this impact can be enhanced or diminished by the students' wider feedback literacy skills.

## Limitations and Recommendations for Future Research

While the findings of this study provide new insights into how feedback literacy and AI literacy shape students' evaluations of GenAI- and teacher-generated feedback, several limitations must be acknowledged, along with recommendations for future research. *First*, the sample was relatively small ( $N = 73$ ) and lacked diversity, as most participants identified as German and female. This imbalance in demographics could restrict the generalizability of the findings to broader student populations in varying cultural or academic contexts. Expanding and diversifying the sample in future research would enhance the external validity of findings. For instance, including students from a range of nationalities, disciplines, and academic levels could produce more generalizable results.

*Second*, the use of self-report measures introduces potential biases, such as social desirability or overestimation of one's competencies (Kreitchmann et al., 2019). Although the adopted instruments demonstrated good reliability and validity, subjective interpretations may still have influenced the results. To gain deeper insights into students' perceptions, future research should combine quantitative methods with qualitative approaches such as interviews, open-ended responses, or think-aloud protocols. These methods could help reveal how students interpret feedback, form trust judgements, and apply critical thinking skills when engaging with GenAI.

*Third*, the study was conducted in an artificial setting where participants evaluated feedback on a pre-written essay rather than feedback on their academic work. This may not fully capture the cognitive and emotional involvement usually linked to actual feedback experiences, which could potentially diminish ecological validity. Evaluating GenAI-generated feedback on students' own assignments could provide a more authentic context for exploring trust, credibility, and engagement.

*Fourth*, the study design was cross-sectional and did not include longitudinal or behavioural measures, such as whether students applied the feedback or revised their work based on it. As a result, it remains uncertain whether positive feedback evaluations led to actual learning behaviours or enhanced academic outcomes. Future studies should incorporate longitudinal data, such as tracking actual revisions, learning outcomes, or feedback uptake over time.

*Fifth*, all participants received feedback from the teacher first, followed by feedback from GenAI. This fixed order may have shaped how students perceived the second source, potentially making it seem repetitive or more credible depending on their expectations. Future research should therefore randomize the order of feedback to reduce systematic bias in students' evaluations.

*Finally*, the study only used quantitative methods. While this allowed for statistical analysis of relationships between variables, it limited the ability to explore students' subjective interpretations, emotional responses, or reasoning processes in greater depth. Moreover, examining additional moderating variables (e.g., prior experience with AI tools, academic motivation, or feedback-seeking behaviour) could enhance understanding of the individual differences that influence feedback engagement.

When combined, these recommendations have the potential to enhance the accuracy of educational assessments and support a more effective integration of GenAI technologies into classroom environments, with a focus on fostering critical engagement and literacy development.

## **Implications for Practice**

The findings from the present study provide valuable insights into how educators can effectively integrate GenAI-generated feedback into learning environments while addressing

students' biases and literacy levels. The results emphasize how crucial it is to promote both feedback literacy and AI literacy to guarantee that students critically engage with feedback generated by AI rather than dismiss it because of presumptions.

One key implication for educators is the necessity of transparency in AI feedback systems. The study demonstrates that students tend to rate GenAI-generated feedback lower once they know its source, stressing the importance of open dialogue about AI's place in learning environments. Educators should openly communicate how GenAI tools function, their strengths and limitations, and how they can complement, rather than replace, traditional feedback mechanisms.

Additionally, enhancing students' feedback literacy should be considered. Since students with higher feedback literacy demonstrated a greater ability to evaluate GenAI-generated feedback based on content rather than source, educators should design course activities that explicitly cultivate feedback literacy. Teaching students how to critically assess the relevance, accuracy, and usefulness of feedback, regardless of whether it comes from a human teacher or GenAI can encourage more effective engagement with GenAI-generated feedback.

Another practical implication is the need to strengthen students' AI literacy so they can effectively engage with and interpret feedback generated by AI systems. The study suggests that students with higher AI literacy are more accepting of GenAI-generated feedback and more capable of evaluating its credibility. To facilitate this, universities should offer clear explanations of how AI models produce responses, how biases may arise in feedback from AI, and how students can fact-check AI-provided information.

Furthermore, combining GenAI-generated feedback with teacher guidance to maximize learning outcomes is advised. GenAI can provide timely, structured feedback, while human teachers can refine and personalize it, addressing relational and emotional aspects that AI systems lack.



Finally, it is essential that institutions support ongoing research and evaluation of GenAI-generated feedback, given its novelty and growing potential. As GenAI continues to evolve, understanding how students engage with its feedback will be key to making the most of its capabilities and embedding it effectively in teaching practices.

By implementing these strategies, educational institutions can optimize the use of GenAI-generated feedback, equipping students with the necessary literacy skills to engage with AI thoughtfully and critically.

## **Conclusion**

In conclusion, this study aimed to investigate how students perceive feedback from both human teachers and GenAI, with a particular focus on the roles of feedback literacy and AI literacy. By examining how these two competencies shape students' evaluations of feedback from different sources, the research sought to explore the cognitive processes underlying feedback perception and uptake.

The findings revealed a significant shift in feedback perception depending on source awareness. When students evaluated feedback blindly, meaning without knowing whether it was generated by GenAI or a human teacher, they rated both feedback pieces similarly. However, upon learning the source, students rated teacher feedback more favourably, while GenAI-generated feedback received lower evaluations. Notably, even students who correctly identified the source beforehand rated teacher feedback more favourably after disclosure.

This suggests an inherent bias favouring human feedback, possibly due to perceptions of credibility, personalization, and instructional depth. Furthermore, feedback literacy was found to be a stronger predictor of positive feedback engagement than AI literacy, particularly influencing how students interacted with GenAI-generated feedback. While AI literacy played

a significant role in fostering trust in feedback from GenAI, feedback literacy was crucial in helping students assess the quality of feedback beyond its source.

Ultimately, this study highlights the importance of fostering both feedback literacy and AI literacy in educational settings. As AI-driven feedback systems become more prevalent, helping students develop the ability to critically assess feedback beyond its source will be essential for fostering informed and effective learning. Strengthening literacy in these areas can facilitate more effective integration of GenAI-generated feedback, enabling students to engage with it meaningfully and productively.

## References

- Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education*, 21(1).  
<https://doi.org/10.1186/s41239-024-00455-4>
- Brown, G. T. L., Peterson, E. R., & Yao, E. S. (2016). Student conceptions of feedback: Impact on self-regulation, self-efficacy, and academic achievement. *British Journal of Educational Psychology*, 86(4), 606–629. <https://doi.org/10.1111/bjep.12126>
- Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, 65(3), 245–281.  
<https://doi.org/10.3102/00346543065003245>
- Casal Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro López, B., & Barro, S. (2023). AI literacy in K-12: a systematic literature review. *International Journal of STEM Education*, 10(1). <https://doi.org/10.1186/s40594-023-00418-7>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325.  
<https://doi.org/10.1080/02602938.2018.1463354>
- Carless, D., & Winstone, N. (2020). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*, 28(1), 150–163.  
<https://doi.org/10.1080/13562517.2020.1782372>
- Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y.-S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, 2(100027), 100027. <https://doi.org/10.1016/j.caeai.2021.100027>

- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 1–18. <https://doi.org/10.1186/s41239-023-00411-8>
- Chiu, T. (2023). The Impact of Generative AI (GenAI) on practices, Policies and Research Direction in education: a Case of ChatGPT and Midjourney. *Interactive Learning Environments*, 32(10), 1–17. <https://doi.org/10.1080/10494820.2023.2253861>
- Corbin, T., Tai, J., & Flenady, G. (2025). Understanding the place and value of GenAI feedback: a recognition-based framework. *Assessment & Evaluation in Higher Education*, 1–14. <https://doi.org/10.1080/02602938.2025.2459641>
- Crompton, H. (2024). *The Promise and Pitfalls of GenAI in Education*. ODU Digital Commons. [https://digitalcommons.odu.edu/stemps\\_fac\\_pubs/369](https://digitalcommons.odu.edu/stemps_fac_pubs/369)
- Dale, S. (2015). Heuristics and biases: The science of decision-making. *Business Information Review*, 32(2), 93–99. <https://doi.org/10.1177/0266382115592536>
- Dawson, P., Yan, Z., Lipnevich, A. A., Tai, J., Boud, D., & Mahoney, P. (2023). Measuring what learners do in feedback: the feedback literacy behaviour scale. *Assessment & Evaluation in Higher Education*, 49(3), 1–15. <https://doi.org/10.1080/02602938.2023.2240983>
- Di Zhang, E., Liu, C., & Yu, S. (2023). The impact of a feedback intervention on university students' second language writing feedback literacy. *Innovations in Education and Teaching International*, 1–17. <https://doi.org/10.1080/14703297.2023.2254275>
- Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00425-2>
- Evans, C. (2013). Making Sense of Assessment Feedback in Higher Education. *Review of Educational Research*, 83(1), 70–120. <https://doi.org/10.3102/0034654312474350>

- Fu, T., & Nassaji, H. (2016). Corrective feedback, learner uptake, and feedback perception in a Chinese as a foreign language classroom. *Studies in Second Language Learning and Teaching*, 6(1), 159. <https://doi.org/10.14746/ssllt.2016.6.1.8>
- Gan, Z., An, Z., & Liu, F. (2021). Teacher feedback practices, student feedback motivation, and feedback behavior: How are they associated with learning outcomes? *Frontiers in Psychology*, 12(1). <https://doi.org/10.3389/fpsyg.2021.697045>
- Grønli, K. M., Walgermo, B. R., McTigue, E. M., & Uppstad, P. H. (2025). Feedback Practices on Young Students' Oral Reading: A Systematic Review. *Review of Educational Research*. <https://doi.org/10.3102/00346543241306070>
- Gul, Dr. R., Tahir, Dr. T., Ishfaq, Dr. U., & Batool, T. (2021). *Impact of Teachers' Workload on Their Time Management Skills at University Level*. Googleusercontent.com. [https://scholar.googleusercontent.com/scholar?q=cache:tjcEkXx5LVoJ:scholar.google.com/+teachers+time+management+skills&hl=de&as\\_sdt=0](https://scholar.googleusercontent.com/scholar?q=cache:tjcEkXx5LVoJ:scholar.google.com/+teachers+time+management+skills&hl=de&as_sdt=0)
- Han, Y., & Xu, Y. (2019). The development of student feedback literacy: the influences of teacher feedback on peer feedback. *Assessment & Evaluation in Higher Education*, 45(5), 1–17. <https://doi.org/10.1080/02602938.2019.1689545>
- Harrison, C. J., Könings, K. D., Schuwirth, L., Wass, V., & van der Vleuten, C. (2014). Barriers to the uptake and use of feedback in the context of summative assessment. *Advances in Health Sciences Education*, 20(1), 229–245. <https://doi.org/10.1007/s10459-014-9524-6>
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Henderson, M., Ryan, T., Boud, D., Dawson, P., Phillips, M., Molloy, E., & Mahoney, P. (2019). The Usefulness of Feedback. *Active Learning in Higher Education*, 22(3). <https://doi.org/10.1177/1469787419872393>

- Heron, M., Medland, E., Winstone, N., & Pitt, E. (2021). Developing the relational in teacher feedback literacy: exploring feedback talk. *Assessment & Evaluation in Higher Education*, 48(2), 172–185. <https://doi.org/10.1080/02602938.2021.1932735>
- Hornberger, M., Bewersdorff, A., & Nerdel, C. (2023). What do university students know about Artificial Intelligence? Development and validation of an AI literacy test. *Computers and Education: Artificial Intelligence*, 5, 100165. <https://doi.org/10.1016/j.caeai.2023.100165>
- Huang, Z., Berry, M. P., Chwyl, C., Hsieh, G., Wei, J., & Forman, E. M. (2025). Comparing Large Language Model AI and Human-Generated Coaching Messages for Behavioral Weight Loss. *Journal of Technology in Behavioral Science*. <https://doi.org/10.1007/s41347-025-00491-5>
- Hutson, J. (2024). Rethinking Plagiarism in the Era of Generative AI. *Journal of Intelligent Communication*, 4(1). <https://doi.org/10.54963/jic.v4i1.220>
- Jensen, L. X., Bearman, M., & Boud, D. (2022). Feedback encounters: towards a framework for analysing and understanding feedback processes. *Assessment & Evaluation in Higher Education*, 1–14. <https://doi.org/10.1080/02602938.2022.2059446>
- Jin, F., Maheshi, B., Lai, W., Li, Y., Gašević, D., Chen, G., Charwat, N., Chan, P. W. K., Martinez-Maldonado, R., Gašević, D., & Tsai, Y.-S. (2025). Students' Perceptions of GenAI-powered Learning Analytics in the Feedback Process: *Journal of Learning Analytics*, 1–17. <https://doi.org/10.18608/jla.2025.8609>
- Kelly, A., Sullivan, M., & Strampel, K. (2023). Generative artificial intelligence: University student awareness, experience, and confidence in use across disciplines. *Journal of University Teaching & Learning Practice*, 20(6). <https://doi.org/10.53761/1.20.6.12>
- Kohnke, L., Zou, D., & Su, F. (2025). Exploring the potential of GenAI for personalised English teaching: Learners' experiences and perceptions. *Computers and Education: Artificial Intelligence*, 8, 100371. <https://doi.org/10.1016/j.caeai.2025.100371>

- Krackov, S. K. (2016). *Giving feedback*. Basicmedical Key.  
<https://basicmedicalkey.com/giving-feedback/>
- Kreitchmann, R. S., Abad, F. J., Ponsoda, V., Nieto, M. D., & Morillo, D. (2019). Controlling for Response Biases in Self-Report Scales: Forced-Choice vs. Psychometric Modeling of Likert Items. *Frontiers in Psychology*, 10.  
<https://doi.org/10.3389/fpsyg.2019.02309>
- Langer, M., König, C. J., Back, C., & Hemsing, V. (2022). Trust in Artificial Intelligence: Comparing Trust Processes between Human and Automated Trustees in Light of Unfair Bias. *Journal of Business and Psychology*, 38. <https://doi.org/10.1007/s10869-022-09829-9>
- Lee, C. C., & Yoke, M. (2024). Using genAI in education: the case for critical thinking. *Frontiers in Artificial Intelligence*, 7. <https://doi.org/10.3389/frai.2024.1452131>
- Lee, S. S., & Moore, R. L. (2024). Harnessing Generative AI (GenAI) for Automated Feedback in Higher Education: A Systematic Review. *Online Learning*, 28(3).  
<https://doi.org/10.24059/olj.v28i3.4593>
- Lee, S., & Song, K. (2024). Teachers' and Students' Perceptions of AI-Generated Concept Explanations: Implications for Integrating Generative AI in Computer Science Education. *Computers and Education Artificial Intelligence*, 7, 100283–100283.  
<https://doi.org/10.1016/j.caeai.2024.100283>
- Lim, S., & Schmälzle, R. (2024). The effect of source disclosure on evaluation of AI-Generated messages: A two-part study. *Computers in Human Behavior Artificial Humans*, 2(1), 100058–100058. <https://doi.org/10.1016/j.chbah.2024.100058>
- Little, T., Dawson, P., Boud, D., & Tai, J. (2023). Can students' feedback literacy be improved? A scoping review of interventions. *Assessment & Evaluation in Higher Education*, 49(1), 1–14. <https://doi.org/10.1080/02602938.2023.2177613>

- Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3313831.3376727>
- Malecka, B., Boud, D., & Carless, D. (2020). Eliciting, processing and enacting feedback: mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education*, 27(7), 1–15. <https://doi.org/10.1080/13562517.2020.1754784>
- Mittal, U., Sai, S., Vinay Chamola, & None Devika. (2024). A Comprehensive Review on Generative AI for Education. *IEEE Access*, 1–1. <https://doi.org/10.1109/access.2024.3468368>
- Nazaretsky, T., Cukurova, M., & Alexandron, G. (2022). An Instrument for Measuring Teachers' Trust in AI-Based Educational Technology. *An Instrument for Measuring Teachers' Trust in AI-Based Educational Technology*. <https://doi.org/10.1145/3506860.3506866>
- Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., & Käser, T. (2024). *AI or Human? Evaluating Student Feedback Perceptions in Higher Education*. <https://doi.org/10.31219/osf.io/6zm83>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Shen, M. Q. (2021). Conceptualizing AI literacy: An Exploratory Review. *Computers and Education: Artificial Intelligence*, 2(1). <https://doi.org/10.1016/j.caeai.2021.100041>
- Rad, H. S., Alipour, R., & Jafarpour, A. (2023). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: a case of Wordtune application. *Interactive Learning Environments*, 32(9), 1–21. <https://doi.org/10.1080/10494820.2023.2208170>
- Raj, M., Berg, J., & Seamans, R. (2023). *Art-ificial Intelligence: The Effect of AI Disclosure on Evaluations of <sup>F1</sup><sub>SEP</sub> Creative Content*. <https://doi.org/10.48550/arxiv.2303.06217>



- Ramu, D., Jain, R., & Jain, A. (2023). *Generation Z's Ability to Discriminate Between AI-generated and Human-Authored Text on Discord*. ArXiv.org.  
<https://doi.org/10.48550/arXiv.2401.04120>
- Raviv, A., Bar-Tal, D., Raviv, A., Biran, B., & Sela, Z. (2003). Teachers' Epistemic Authority: Perceptions of Students and Teachers. *Social Psychology of Education*, 6(1), 17–42. <https://doi.org/10.1023/a:1021724727505>
- Rawal, S. B. (2023). Exploring Students' Perceptions on Teachers' Feedback in the Classroom (A Study of Achham Multiple Campus, Nepal). *World Journal of Education and Humanities*, 5(4), p58–p58. <https://doi.org/10.22158/wjeh.v5n4p58>
- Richmond, J. L., & Nicholls, K. (2024). Using Generative AI to Promote Psychological, Feedback, and Artificial Intelligence Literacies in Undergraduate Psychology. *Teaching of Psychology*. <https://doi.org/10.1177/00986283241287203>
- Roe, J., Perkins, M., & Ruelle, D. (2024). *Understanding student and academic staff perceptions of AI use in assessment and feedback*. Intelligent Technologies in Education; Advanced Online Publication. <https://doi.org/10.70770/rzzz6y35>
- Rowe, A. D., & Wood, L. N. (2009). Student Perceptions and Preferences for Feedback. *Asian Social Science*, 4(3). <https://doi.org/10.5539/ass.v4n3p78>
- Ruegg, R. (2015). Differences in the Uptake of Peer and Teacher Feedback. *RELC Journal*, 46(2), 131–145. <https://doi.org/10.1177/0033688214562799>
- Saihi, A., Ben-Daya, M., Hariga, M., & Asad, R. (2024). A Structural Equation Modeling Analysis of Generative AI Chatbots Adoption Among Students and Educators in Higher Education. *Computers and Education Artificial Intelligence*, 100274–100274. <https://doi.org/10.1016/j.caeai.2024.100274>
- Shailendra, S., Kadel, R., & Sharma, A. (2024). Framework for Adoption of Generative Artificial Intelligence (GenAI) in Education. *IEEE Transactions on Education*, 1–9. <https://doi.org/10.1109/te.2024.3432101>

- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Strijbos, J.-W., Pat-El, R., & Narciss, S. (2021). Structural validity and invariance of the Feedback Perceptions Questionnaire. *Studies in Educational Evaluation*, 68, 100980. <https://doi.org/10.1016/j.stueduc.2021.100980>
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a review. *Assessment & Evaluation in Higher Education*, 30(4), 325–341. <https://doi.org/10.1080/02602930500099102>
- Tanes, Z., Arnold, K. E., King, A. S., & Remnet, M. A. (2011). Using Signals for appropriate feedback: Perceptions and practices. *Computers & Education*, 57(4), 2414–2422. <https://doi.org/10.1016/j.compedu.2011.05.016>
- Tubino, L., & Adachi, C. (2022). Developing feedback literacy capabilities through an AI automated feedback tool. *ASCILITE Publications*, e22039. <https://doi.org/10.14742/apubs.2022.39>
- Tzirides, A. O. (Olnancy), Zapata, G., Kastania, N. P., Saini, A. K., Castro, V., Ismael, S. A., You, Y., Santos, T. A. dos, Sears Smith, D., O'Brien, C., Cope, B., & Kalantzis, M. (2024). Combining human and artificial intelligence for enhanced AI literacy in higher education. *Computers and Education Open*, 6, 100184. <https://doi.org/10.1016/j.caeo.2024.100184>
- van de Ridder, J. M. M., Berk, F. C. J., Stokking, K. M., & ten Cate, O. Th. J. (2014). Feedback providers' credibility impacts students' satisfaction with feedback and delayed performance. *Medical Teacher*, 37(8), 767–774. <https://doi.org/10.3109/0142159x.2014.970617>
- Van der Kleij, F. M., & Lipnevich, A. A. (2020). Student perceptions of assessment feedback: a critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability*, 33. <https://doi.org/10.1007/s11092-020-09331-x>

- Vora, V., et al. (2023). A Multimodal Approach for Detecting AI Generated Content using BERT and CNN. *International Journal on Recent and Innovation Trends in Computing and Communication*, 11(9), 691–701.  
<https://doi.org/10.17762/ijritcc.v11i9.8861>
- Wang, C., Wang, H., Li, Y., Dai, J., Gu, X., & Yu, T. (2024). Factors Influencing University Students' Behavioral Intention to Use Generative Artificial Intelligence: Integrating the Theory of Planned Behavior and AI Literacy. *International Journal of Human-Computer Interaction*, 1–23. <https://doi.org/10.1080/10447318.2024.2383033>
- Wei, W., Sun, Y., & Xu, X. (2020). Investigating the impact of increased student feedback literacy level on their expectations on university teachers' feedback. *Assessment & Evaluation in Higher Education*, 1–12.  
<https://doi.org/10.1080/02602938.2020.1846017>
- Wetzler, E. L., Cassidy, K. S., Jones, M. J., Frazier, C. R., Korbut, N. A., Sims, C. M., Bowen, S. S., & Wood, M. (2024). Grading the Graders: Comparing Generative AI and Human Assessment in Essay Evaluation. *Teaching of Psychology*.  
<https://doi.org/10.1177/00986283241282696>
- Winstone, N. E., Mathlin, G., & Nash, R. A. (2019). Building Feedback Literacy: Students' Perceptions of the Developing Engagement With Feedback Toolkit. *Frontiers in Education*, 4. <https://doi.org/10.3389/feduc.2019.00039>
- Woodrow, J., Malik, A., & Piech, C. (2024). AI Teaches the Art of Elegant Coding: Timely, Fair, and Helpful Style Feedback in a Global Course. *SIGCSE 2024: The 55th ACM Technical Symposium on Computer Science Education*, 1442–1448.  
<https://doi.org/10.1145/3626252.3630773>
- Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. *ArXiv (Cornell University)*.  
<https://doi.org/10.48550/arxiv.2408.12143>

- Zacharias, N. T. (2007). Teacher and Student Attitudes toward Teacher Feedback. *RELC Journal*, 38(1), 38–52. <https://doi.org/10.1177/0033688206076157>
- Zaman, B. U. (2023). Transforming Education Through AI, Benefits, Risks, and Ethical Considerations. *Transforming Education through AI, Benefits, Risks, and Ethical Considerations*. <https://doi.org/10.36227/techrxiv.24231583.v1>
- Zhan, Y., Boud, D., Dawson, P., & Yan, Z. (2025). Generative artificial intelligence as an enabler of student feedback engagement: a framework. *Higher Education Research & Development*, 1–16. <https://doi.org/10.1080/07294360.2025.2476513>
- Zhan, Y., & Yan, Z. (2025). Students' engagement with ChatGPT feedback: implications for student feedback literacy in the context of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 1–14. <https://doi.org/10.1080/02602938.2025.2471821>
- Zhang, Y., Schunn, C. D., & Wu, Y. (2024). Interconnecting peer feedback literacy: Exploring the relationship between providing and acting on peer feedback. *Studies in Educational Evaluation*, 83, 101411–101411. <https://doi.org/10.1016/j.stueduc.2024.101411>

## **Appendices**

### **Appendix A**

#### *AI statement*

For writing this paper, the author used ChatGPT and Copilot to assist with R coding and the interpretation of statistical outputs. Moreover, Grammarly was used to check for spelling as well as grammar errors. Consensus was utilized to support and refine the literature search process. All content was subsequently reviewed, edited, and finalized by the author, who takes full responsibility for the work.

## Appendix B

### *Consent Form for “Generative AI or Human? Exploring the Role of AI- and Feedback*

#### *Literacy in Student Feedback Perceptions”*

**Please tick the appropriate boxes**

**Yes No**

#### **Taking part in the study**

I have read and understood the study information dated [DD/MM/YYYY], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.

☐ ☐

I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions, and I can withdraw from the study at any time, without having to give a reason.

☐ ☐

I understand that taking part in the study involves two parts, first writing an online essay of max. 300 words and then evaluating two feedback pieces for this assignment in an online survey questionnaire that you can fill in yourself in the second part. Both sessions will each take approximately 20-30 minutes.

☐ ☐

#### **Use of the information in the study**

I understand that information I provide will be used for writing a Bachelor thesis report that might be published.

☐ ☐

I understand that personal information collected about me that can identify me, such as my name, my profession, my age or where I live, will not be shared beyond the study team.

☐ ☐

I agree that my information can be quoted in research outputs.

☐ ☐

#### **Future use and reuse of the information by others**

I give permission for the survey database that I provide to be archived in the University of Twente Student Theses database, so it can be used for future research and learning.

☐ ☐

#### **Signatures**

\_\_\_\_\_  
Name of participant [printed]

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Date

**Study contact details for further information:**

Researcher:

Supervisor:

**Contact Information for Questions about Your Rights as a Research Participant**

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee/domain Humanities & Social Sciences of the Faculty of Behavioural, Management and Social Sciences at the University of Twente

## Appendix C

### *Feedback Literacy Scale*

For the next series of statements, please read carefully and indicate how much you agree or disagree on the scale of 1 to 6!

Answer Options to each of the following items: Never, Almost Never, Rarely, Sometimes, Almost Always, Always

1. I reflect on the quality of my own work and use my reflection as a source of information to improve my work.
2. I seek out examples of good work to improve my work.
3. When other people provide me with input about my work I listen or read thoughtfully.
4. When I am working on a task, I consider comments I have received on similar tasks.
5. I ask for comments about specific aspects of my work.
6. I carefully consider comments about my work before deciding if I will use them or not.
7. When receiving conflicting information about my work from different sources, I make a judgement about what I will use.
8. When deciding what to do with comments, I consider the credibility of their sources.
9. I consider how comments relate to criteria or standards.
10. I check whether my work is better after I have acted on comments.
11. I use comments on my work to refine my understanding of what good quality work looks like.
12. When receiving comments I plan how I will use them to improve my future work, not just the immediate task.
13. I keep comments on previous work to use again in the future.
14. When I receive comments from others, I use them to improve what I'm working on at the time.
15. When commenting on the work of others, I provide constructive criticism.
16. I comment on other people's work when I am invited to.
17. When commenting on other people's work I refer to standards or criteria.
18. I offer to provide feedback to my peers.
19. I try to be very clear when providing feedback comments to others.
20. I am open to reasonable criticism about my work.
21. I deal well with any negative emotional responses I have to feedback information.



22. When a feedback message is valuable but upsetting or annoying, I still find a way to make use of it.
23. Critical comments motivate me to improve my work.
24. I make use of critical comments even if they are difficult to receive.

## Appendix D

### *AI Literacy Scale*

For the next series of statements, please read carefully and indicate how much you agree or disagree on the scale of 1 to 7!

(Answer Options to each of the following statements: Strongly Disagree, Disagree, Somewhat Disagree, Neutral, Somewhat Agree, Agree, Strongly Agree)

1. I understand how GAI (Generative Artificial Intelligence) products like ChatGPT achieve human-machine interaction.
2. I can recognize the artificial intelligence technologies used in the applications (e.g. Douyin, Taobao) and products (e.g. robotic vacuum cleaners) that I use.
3. I understand why GAI technology relies on big data.
4. I understand how GAI technology optimizes the translation output of online translation.
5. I understand how GAI products process images to achieve visual recognition functionality.
6. I know how GAI products like ChatGPT and Gemini perform speech recognition tasks.
7. I can proficiently use AI applications (e.g. ChatGPT or Sora) or products to assist me in daily tasks.
8. I usually find it easy to learn how to use new GAI applications or products.
9. I can use GAI applications or products to enhance my work efficiently.
10. I can evaluate the functionality and limitations of GAI applications (e.g. ChatGPT or Midjourney) or products after using them for a period.
11. I can choose the appropriate solution from the various solutions provided by GAI-related applications and products.
12. I can select the most suitable GAI application or product for various specific tasks.
13. When using GAI applications or products, I always adhere to ethical principles.
14. When using GAI applications or products, I am vigilant about privacy and information security issues.
15. I am always vigilant about the misuse of GAI technology.
16. My parents support me in learning how to use GAI technology.
17. Most people I know believe that I should learn how to use GAI technology.
18. My classmates believe it is necessary to learn how to use GAI technology.
19. My teachers believe it is necessary to learn how to use GAI technology.

20. Learning GAI-related technologies is relatively easy for me.
21. Applying GAI technology to assist me in work and study is relatively easy for me.
22. Applying GAI technology to solve problems in daily life is relatively easy for me.
23. I think it is very wise to apply GAI technology to solve problems in daily life.
24. I find using GAI technology enjoyable.
25. I find using GAI technology useful for my learning and work.
26. I intend to continue using GAI technology in the future.
27. GAI technology has been very helpful for learning, so I will continue using it.
28. I am willing to introduce GAI technology to my peers.



## Appendix E

### *The Essay*

**The following argumentative essay has been written by a student. Please read it carefully before proceeding to the next step.**

#### **Should university students use Generative AI tools in their studies?**

The topic of whether university students should be allowed to use generative AI tools in their studies is very recent and controversial. While critics argue that AI threatens academic integrity, others recognize its potential to enhance learning. Instead of simply forbidding students to use generative AI (GenAI), universities should see it as a valuable educational tool. However, ethical guidelines need to be implemented to make sure the tool is used responsibly.

There are many advantages to using GenAI. Firstly, AI tools, such as ChatGPT or Grammarly, have the potential to enhance learning and efficiency. They can help with brainstorming ideas, refining work, and also summarizing and explaining complex materials. Instead of spending a large amount of time on less important things like finding the correct wording and ensuring correct spelling, students can focus on more important things like gaining knowledge through literature reviews or thinking of innovative arguments. Furthermore, AI is easily accessible for everybody and available 24/7, so the students receive explanations and help immediately, even outside of scheduled classes. This makes AI an efficient help for students and can motivate them as well.

Apart from that, GenAI also promotes accessibility and inclusivity. Therefore, it can be beneficial for students who are non-native speakers or who have learning disabilities, and also for students who do not have much time due to family or work. GenAI can translate material into various languages and can offer personalized responses to them. ChatGPT, for example, creates a human-like dialogue with the user, and whenever something is unclear, it can refine its answer and try to explain the material differently. This adaptability ensures that students receive support that fits their individual learning needs.

Lastly, it is crucial to mention that in today's technology-driven world, being able to understand and use AI efficiently (AI literacy) is becoming an essential skill for many workplaces. AI is being integrated into daily tasks, from scheduling and research assistance to data analysis and problem solving. Universities, therefore, have a responsibility to prepare students for an AI-driven job market by introducing them to GenAI and teaching them how to use it effectively.

Although critics argue that AI promotes plagiarism and discourages independent thinking, simply ignoring it is not the right solution. Instead, universities should create clear guidelines so that AI does not do all the thinking and work for students. Moreover, they should integrate digital literacy training into classes so that students have the skills to use GenAI responsibly. Thereby, GenAI can become a powerful tool that enhances education.

## Appendix F

### *Feedback 2: GenAI-generated Feedback*

**The student has received another feedback from an unknown source again. Please read it carefully and rate its quality using the following rubric.**

- Your essay clearly introduces the topic of GenAI and highlights the controversy surrounding its use in academic studies. Your position that GenAI should be permitted under clear guidelines is explicit and persuasive.
- Well done in establishing a definite stance early in your essay. You offer several compelling reasons supporting your viewpoint, such as increased efficiency, accessibility, inclusivity, and preparation for the AI-driven job market. Your use of specific examples like Grammarly and ChatGPT effectively supports your claims.
- However, to further strengthen your reasoning, consider incorporating more varied and concrete evidence, such as recent statistics, articles, or expert opinions about GenAI's impact on academic performance.
- Regarding counterarguments, you thoughtfully acknowledge concerns about plagiarism and reduced independent thinking. Your suggestion to implement clear guidelines and digital literacy training is an effective refutation, but enhancing this section with specific examples of successful implementations or expert recommendations would strengthen your argument further.
- Your conclusion effectively summarizes your key points and reinforces your main thesis, emphasizing GenAI's potential when responsibly integrated.
- Overall, the essay is coherent and persuasive, though adding more

diverse evidence and expanding on your refutation of counterarguments could improve its depth and credibility further.



## Appendix G

### *Feedback Prompt for ChatGPT*

You are a professor of academic writing at the undergraduate level. Your student has submitted an argumentative essay as an assignment for this course on the role of artificial intelligence in education, which should be no more than 500 words in length. You now need to provide the student with feedback on the quality of their argumentative writing, identifying problems in their essay and suggesting solutions. Your feedback should be given in a paragraph containing 250 to 300 words and address the following questions:

- Has the student provided an introduction relevant to the topic?
- Has the student presented a clear and definite position regarding the topic?
- Has the student provided reasons to support their position? And have they substantiated these reasons with credible evidence (in the form of examples, personal experiences, statistics, expert opinions, and research evidence)?
- Has the student presented any counterarguments to their position? And have they refuted these counterarguments with credible evidence (in the form of examples, personal experiences, statistics, expert opinions, and research evidence)?
- Has the student concluded their essay effectively?

Please use a step-by-step approach to respond to this request.

-----  
For example, a student has sent you the following argumentative essay:

“The role of artificial intelligence in education

Today, artificial intelligence is used extensively in all fields, especially in education. The use of AI tools in education has benefits such as creating smart educational content, learning in any language, creating tests, etc., but alongside such benefits are challenges and limitations such as lack of quality data, cost, and access. Generally, the benefits of using AI tools outweigh the challenges, and in my opinion, these tools can be used in the classroom to great advantage because firstly, using these tools improves learners' education—for example, by making many of the abstract contents of courses like physics and chemistry concrete, learning occurs faster and better. Secondly, teachers can receive ideas and help regarding their teaching methods from these tools and clarify educational material for students. For example, teachers for lower grades can use games and entertainments provided by AI tools. Thirdly, with constant access to artificial intelligence, teachers and learners can use it at any time to get more information and enhance their knowledge. Students can easily have essays and any needed texts in class. Teachers can also create34 creative exercises and smart tests through these tools and even use AI for grading and checking tests. Therefore, artificial intelligence can create a major revolution in the education industry and cause an increase in productivity in various aspects of it.”

-----  
Now you want to provide feedback to improve this argumentative text. To prepare this feedback, you should act as follows. First, check whether the student has provided a suitable introduction before stating their own position. If the introduction is weak or incomplete, offer suggestions for its improvement. Next, check whether the student has clearly and explicitly stated their position. If the position is vague, offer suggestions for its improvement. Then, it should be checked whether the student has provided valid evidence (in the form of examples, personal experience, statistics, expert opinion, and research evidence) in support of their position. If the evidence is insufficient or not credible, offer suggestions for its improvement. Next, it should be noted whether the student has raised opposing views to their position and refuted them with valid evidence (in the form of examples, personal experience, statistics, expert opinion, and research evidence). If this part is weak, offer

suggestions for its improvement. Finally, check whether the student has provided an extensive summary. If the summary is insufficient, offer suggestions for its improvement. Following these considerations, a sample 250 to 300-word feedback for the argumentative text above could be as follows:

"In the beginning of the text (the first four lines) you mention the advantages and challenges and at the end you also state your position in the form 'In my opinion, these tools can be used in the classroom to great advantage.' It would be better instead of stating these advantages and disadvantages to start your text with a brief introduction that clarifies the background of the topic and creates an engaging entry into the subject. After this introduction, you should state your position clearly.

After stating your position, you have discussed one reason, saying that the use of these tools makes many of the abstract contents of courses concrete and then mentioned several courses like physics and chemistry as examples or evidence, which is excellent. In the second reason, you have stated that teachers can use artificial intelligence in their teaching methods. Then you have mentioned an example or evidence that games and entertainments produced by AI tools can be used. As the third reason, you talk about teachers having access to artificial intelligence at any time which can update their knowledge, but for this reason, you have not mentioned evidence in the form of examples, expert opinions, personal experiences, and statistics.

Finally, you have clearly stated your conclusion. However, it would have been better to mention the negative aspects of using AI tools in the classroom as well, after referring to the positive aspects. Providing evidence to refute these negative aspects and then stating conclusions based on that would strengthen your argument, making it easier for the reader to accept."

## Appendix H

### *Feedback 1: Teacher Feedback*

**The student has received the feedback below from an unknown source. Please read the feedback carefully and evaluate its quality using the following rubrics!**

- When you write an essay on a controversial topic, it is always good to take a clear position in the introductory paragraph of your essay. While I understand that you are in favor of using GenAI, I believe this clarification increases the transparency of your essay for the readers.
- I consider this essay an academic writing assignment, and it is always important to provide evidence and references for the claims you make. Therefore, I suggest adding relevant references to different parts of your essay.
- I believe this essay can be more convincing if you add examples. In your essays, in different parts, you mentioned different potentials and benefits of GenAI in education, such as enhancing learning, brainstorming, summarizing, etc. However, you did not give an example to clarify this. How can GenAI do that?
- Another note is to have counterarguments and responses to counterarguments. It is good to see that in your essay, in the beginning and the last paragraph, you highlighted counterarguments such as risks for promoting plagiarism, threatening independent thinking, and academic integrity. However, you did not provide any responses to those critics. Are those criticisms valid, or is there a way to make sure that GenAI is used in an ethical, responsible, and effective way?
- Last but not least, I believe a good conclusion is a necessity in the end to have a good wrapping up.

## **Appendix I**

### *Feedback Evaluation Form*

To what extent do you associate this feedback/feedback provider with the following term?

Answer Options: Strongly Disagree, Somewhat Disagree, Neither Agree nor Disagree,  
Somewhat Agree, Strongly Agree

1. Precise
2. Fair
3. Factual
4. Relevant
5. Informative
6. Applicable
7. Authentic
8. Sincere

## Appendix J

### *R Script*

```
rm(list = ls())

library(broom)
library(tidyverse)
library(janitor)
library(readxl)
library(ggplot2)
library(interactions)
library(modelr)
library(stats)
library(tidyr)
library(dplyr)
library(ez)
library(stringr)
library(mediation)
library(bda)
library(afex)
library(car)

#importing Dataset: Data2
View(Data2)
#Descriptive analysis
Data2 %>%
  ggplot() +
  geom_bar(aes(x = Age))
Data2 %>%
  summarise(mean = mean(Age), sd = sd(Age), var = var(Age), minimum = min(Age),
            maximum = max(Age))
summary(Data2$Age)

###Because the mean was not computed
Data2 <- Data2 %>%
  mutate(Age = as.numeric(as.character(Age)))

Data2 %>%
  tabyl(Gender)
Data2 %>%
  tabyl(Gender) %>%
  adorn_percentages("col") %>%
  adorn_pct_formatting(digits = 1)
```

```

Data2 %>%
  tabyl(Nationality)
Data2 %>%
  tabyl(Nationality) %>%
  adorn_percentages("col") %>%
  adorn_pct_formatting(digits = 1)

#Feedback & AI literacy Score M & SD
FLM_mean_value <- mean(Data2$`F-LSM`, na.rm = TRUE)
FLM_sd_value <- sd(Data2$`F-LSM`, na.rm = TRUE)

ALM_mean_value <- mean(Data2$`AI-LSM`, na.rm = TRUE)
ALM_sd_value <- sd(Data2$`AI-LSM`, na.rm = TRUE)

#Evaluation scores M & SD for both feedbacks before & after reveal
F1B_mean_value <- mean(Data2$`F1B`, na.rm = TRUE)
F1B_sd_value <- sd(Data2$`F1B`, na.rm = TRUE)

F2B_mean_value <- mean(Data2$`F2B`, na.rm = TRUE)
F2B_sd_value <- sd(Data2$`F2B`, na.rm = TRUE)

F1A_mean_value <- mean(Data2$`F1A`, na.rm = TRUE)
F1A_sd_value <- sd(Data2$`F1A`, na.rm = TRUE)

F2A_mean_value <- mean(Data2$`F2A`, na.rm = TRUE)
F2A_sd_value <- sd(Data2$`F2A`, na.rm = TRUE)

#RQ1
##Keep only relevant columns and rename them
feedback_data <- data.frame(
  Teacher_Blind = Data2$`F1B`,
  AI_Blind = Data2$`F2B`,
  Teacher_Known = Data2$`F1A`,
  AI_Known = Data2$`F2A`
)

feedback_long <- feedback_data %>%
  mutate(Participant = row_number()) %>%
  pivot_longer(
    cols = -Participant,
    names_to = c("Source", "Awareness"),
    names_sep = "_",
    values_to = "Rating"
  )

```

```

## Prepare data for ezANOVA
anova_data <- feedback_long %>%
  mutate(Source = factor(Source),
         Awareness = factor(Awareness),
         Participant = factor(Participant),
         AI_LSM = Data2$`AI-LSM`[Participant],
         F_LSM = Data2$`F-LSM`[Participant])

ezANOVA(
  data = anova_data,
  dv = Rating,
  wid = Participant,
  within = .(Source, Awareness),
  detailed = TRUE
)

# Use simplified model for assumption checks only
aov_model_diag <- aov(Rating ~ Source * Awareness, data = anova_data)

# Plot diagnostics
par(mfrow = c(2, 2))
plot(aov_model_diag)

# Normality test on residuals
shapiro.test(residuals(aov_model_diag))

ggplot(anova_data, aes(x = Awareness, y = Rating, fill = Source)) +
  stat_summary(fun = mean, geom = "bar", position = "dodge") +
  stat_summary(fun.data = mean_se, geom = "errorbar",
              position = position_dodge(width = 0.9), width = 0.2) +
  labs(title = "Perceived Quality of Feedback by Source and Awareness",
       y = "Mean Rating", x = "Source Awareness") +
  theme_minimal()

#Part 2 of RQ1: Does the Source Effect Hold for Correct Guesses Only?
# Step 1: Add variables indicating whether the guess was correct
Data2 <- Data2 %>%
  mutate(
    Correct_Teacher = `Source Guess F1` == "Teacher",
    Correct_GenAI = `Source Guess F2` == "AI"
  )

# Step 2: Filter only participants who correctly identified both sources

```

```

correct_guess_data <- Data2 %>%
  filter(Correct_Teacher == TRUE, Correct_GenAI == TRUE)

# Step 3: Compare evaluations of known teacher vs. GenAI feedback in this subgroup
t.test(correct_guess_data$F1A, correct_guess_data$F2A, paired = TRUE)

# Step 4: Descriptive means
mean(correct_guess_data$F1A, na.rm = TRUE)
mean(correct_guess_data$F2A, na.rm = TRUE)

#for participants who correctly guessed teacher feedback (regardless of GenAI guess)
teacher_correct <- Data2 %>% filter(Correct_Teacher == TRUE)

# Paired t-test comparing F1A (teacher feedback rating) and F2A (GenAI feedback rating)
t.test(teacher_correct$F1A, teacher_correct$F2A, paired = TRUE)

# Descriptive means
mean(teacher_correct$F1A, na.rm = TRUE)
mean(teacher_correct$F2A, na.rm = TRUE)

#for participants who correctly guessed GenAI feedback (regardless of teacher guess)
genai_correct <- Data2 %>% filter(Correct_GenAI == TRUE)

# Paired t-test comparing F1A (teacher feedback rating) and F2A (GenAI feedback rating)
t.test(genai_correct$F1A, genai_correct$F2A, paired = TRUE)

# Descriptive means
mean(genai_correct$F1A, na.rm = TRUE)
mean(genai_correct$F2A, na.rm = TRUE)

#Standard errors
# For participants who guessed BOTH correctly
both_correct <- Data2 %>% filter(Correct_Teacher == TRUE, Correct_GenAI == TRUE)

both_summary <- both_correct %>%
  summarise(
    n = n(),
    mean_teacher = mean(F1A, na.rm = TRUE),
    sd_teacher = sd(F1A, na.rm = TRUE),
    se_teacher = sd_teacher / sqrt(n),
    mean_genai = mean(F2A, na.rm = TRUE),
    sd_genai = sd(F2A, na.rm = TRUE),
    se_genai = sd_genai / sqrt(n)
  )

```



```
# For participants who guessed Teacher correctly (regardless of GenAI)
teacher_correct <- Data2 %>% filter(Correct_Teacher == TRUE)
```

```
teacher_summary <- teacher_correct %>%
  summarise(
    n = n(),
    mean_teacher = mean(F1A, na.rm = TRUE),
    sd_teacher = sd(F1A, na.rm = TRUE),
    se_teacher = sd_teacher / sqrt(n),
    mean_genai = mean(F2A, na.rm = TRUE),
    sd_genai = sd(F2A, na.rm = TRUE),
    se_genai = sd_genai / sqrt(n)
  )
```

```
# For participants who guessed GenAI correctly (regardless of Teacher)
genai_correct <- Data2 %>% filter(Correct_GenAI == TRUE)
```

```
genai_summary <- genai_correct %>%
  summarise(
    n = n(),
    mean_teacher = mean(F1A, na.rm = TRUE),
    sd_teacher = sd(F1A, na.rm = TRUE),
    se_teacher = sd_teacher / sqrt(n),
    mean_genai = mean(F2A, na.rm = TRUE),
    sd_genai = sd(F2A, na.rm = TRUE),
    se_genai = sd_genai / sqrt(n)
  )
```

```
# Print all summaries
```

```
both_summary
teacher_summary
genai_summary
```

```
# For both_correct group
sd_genai_both <- sd(both_correct$F2A, na.rm = TRUE)
n_both <- nrow(both_correct)
se_genai_both <- sd_genai_both / sqrt(n_both)
se_genai_both
```

```
# For teacher_correct group
sd_genai_teacher <- sd(teacher_correct$F2A, na.rm = TRUE)
n_teacher <- nrow(teacher_correct)
se_genai_teacher <- sd_genai_teacher / sqrt(n_teacher)
```

```
se_genai_teacher
```

```
# For genai_correct group
```

```
sd_genai_genai <- sd(genai_correct$F2A, na.rm = TRUE)
```

```
n_genai <- nrow(genai_correct)
```

```
se_genai_genai <- sd_genai_genai / sqrt(n_genai)
```

```
se_genai_genai
```

```
summary_df <- data.frame(
```

```
  Group = rep(c("Both Correct (n=20)", "Teacher Correct (n=35)", "GenAI Correct (n=26)"),  
  each = 2),
```

```
  Feedback = rep(c("Teacher", "GenAI"), times = 3),
```

```
  Mean = c(4.38, 3.66, 4.41, 3.72, 4.27, 3.67),
```

```
  SE = c(0.111, 0.195, 0.080, 0.135, 0.114, 0.187)
```

```
)
```

```
ggplot(summary_df, aes(x = Group, y = Mean, fill = Feedback)) +
```

```
  geom_bar(stat = "identity", position = position_dodge(width = 0.7), width = 0.6) +
```

```
  geom_errorbar(aes(ymin = Mean - SE, ymax = Mean + SE),
```

```
    position = position_dodge(width = 0.7), width = 0.25) +
```

```
  labs(title = "Feedback Ratings by Source Guess Accuracy",
```

```
    y = "Mean Feedback Rating",
```

```
    x = "Participant Group") +
```

```
  scale_fill_manual(values = c("Teacher" = "#1f77b4", "GenAI" = "#ff7f0e")) +
```

```
  theme_minimal() +
```

```
  theme(
```

```
    text = element_text(size = 12),
```

```
    legend.title = element_blank()
```

```
) +
```

```
  geom_text(
```

```
    aes(label = round(Mean, 2)),
```

```
    position = position_dodge(width = 0.7),
```

```
    vjust = -0.8,
```

```
    size = 3.5
```

```
) +
```

```
  ylim(0, 5)
```

```
# Summary data frame (already created earlier)
```

```
summary_df <- data.frame(
```

```
  Group = rep(c("Both Correct (n=20)", "Teacher Correct (n=35)", "GenAI Correct (n=26)"),  
  each = 2),
```

```
  Feedback = rep(c("Teacher", "GenAI"), times = 3),
```

```
  Mean = c(4.38, 3.66, 4.41, 3.72, 4.27, 3.67),
```

```
  SE = c(0.111, 0.195, 0.080, 0.135, 0.114, 0.187)
```

)

```
# Ensure correct factor ordering
summary_df$Group <- factor(summary_df$Group,
                           levels = c("Both Correct (n=20)", "Teacher Correct (n=35)", "GenAI Correct
(n=26)"))
```

```
# Final polished plot
```

```
ggplot(summary_df, aes(x = Group, y = Mean, fill = Feedback)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.7), width = 0.6) +
  geom_errorbar(aes(ymin = Mean - SE, ymax = Mean + SE),
               position = position_dodge(width = 0.7), width = 0.25) +
  geom_text(aes(label = round(Mean, 2)),
           position = position_dodge(width = 0.7), vjust = -0.8, size = 3.5) +
  scale_fill_manual(values = c("Teacher" = "#1f77b4", "GenAI" = "#ff7f0e")) +
  ylim(0, 5) + # Adjust as needed for label clearance
  labs(title = "Feedback Ratings by Source Guess Accuracy",
       y = "Mean Feedback Rating",
       x = "Participant Group") +
  theme_minimal() +
  theme(
    text = element_text(size = 12),
    panel.grid.major.y = element_line(color = "gray85"),
    legend.title = element_blank()
  )
```

```
#Part 3 of RQ1
```

```
#Correct & incorrect source identifications
```

```
## True sources
```

```
true_sources <- data.frame(
  Feedback = c("Teacher", "AI"),
  True_Source = c("Teacher", "AI")
)
```

```
##Long format
```

```
source_id <- data.frame(
  Participant = 1:nrow(Data2),
  Feedback = rep(c("Teacher", "AI"), each = nrow(Data2)),
  Selected_Source = c(Data2$`Source Guess F1`, Data2$`Source Guess F2`)
)
```

```
##True sources to determine correctness
```

```
source_id <- source_id %>%
  left_join(true_sources, by = "Feedback") %>%
```

```

mutate(Accuracy = ifelse(Selected_Source == True_Source, "Correct", "Incorrect"))

ggplot(source_id, aes(x = Feedback, fill = Selected_Source)) +
  geom_bar(position = "dodge") +
  labs(title = "Participant Identification of Feedback Source",
        x = "Feedback Type", y = "Number of Selections",
        fill = "Selected Source") +
  theme_minimal()

ggplot(source_id, aes(x = Feedback, fill = Accuracy)) +
  geom_bar(position = "dodge") +
  labs(title = "Accuracy of Source Identification",
        x = "Feedback Type", y = "Number of Participants",
        fill = "Accuracy") +
  scale_fill_manual(values = c("Correct" = "#4CAF50", "Incorrect" = "#F44336")) +
  theme_minimal()

##Counts of each source selection
table(source_id$Feedback, source_id$Selected_Source)

##Accuracy counts
table(source_id$Feedback, source_id$Accuracy)

#RQ2
# Ensure No Missing Values Before Analysis
Data2 <- Data2 %>% drop_na(`F-LSM`, `AI-LSM`, `F2A`, `F2B`, `F1A`, `F1B`)
feedback_data <- feedback_data %>% drop_na(Teacher_Blind, AI_Blind, Teacher_Known,
AI_Known)
feedback_long <- feedback_long %>% drop_na(Participant, Source, Awareness, Rating)

# Standardize Feedback Literacy (Scale: 1-6) & AI Literacy (Scale: 1-7)
Data2 <- Data2 %>%
  mutate(F_LSM_std = scale(`F-LSM`),
         AI_LSM_std = scale(`AI-LSM`))

# 1. Correlation Analysis: Feedback Literacy vs Feedback Evaluation (GenAI & Teacher)
cor.test(Data2$F_LSM_std, Data2$F2B, use = "complete.obs")
cor.test(Data2$F_LSM_std, Data2$F2A, use = "complete.obs")
cor.test(Data2$F_LSM_std, Data2$F1B, use = "complete.obs")
cor.test(Data2$F_LSM_std, Data2$F1A, use = "complete.obs")

#Teacher vs AI (Blind)
cor.test(feedback_data$Teacher_Blind, feedback_data$AI_Blind, use = "complete.obs")
# Teacher vs AI (Known)

```

```
cor.test(feedback_data$Teacher_Known, feedback_data$AI_Known, use = "complete.obs")
```

```
# 2.Linear Regression: Feedback Literacy Predicting Feedback Evaluation (GenAI & Teacher)
```

```
lm_feedback_genai <- lm(F2A ~ F_LSM_std, data = Data2)
summary(lm_feedback_genai)
```

```
#Assumption check
```

```
lm_feedback_genai <- lm(F2A ~ F_LSM_std, data = Data2)
par(mfrow = c(2, 2))
plot(lm_feedback_genai)
shapiro.test(residuals(lm_feedback_genai))
ncvTest(lm_feedback_genai)
par(mfrow = c(1, 1))
qqPlot(lm_feedback_genai)
```

```
lm_feedback_teacher <- lm(F1A ~ F_LSM_std, data = Data2)
summary(lm_feedback_teacher)
```

```
#Assumption check
```

```
lm_feedback_teacher <- lm(F1A ~ F_LSM_std, data = Data2)
par(mfrow = c(2, 2))
plot(lm_feedback_teacher)
shapiro.test(residuals(lm_feedback_teacher))
ncvTest(lm_feedback_teacher)
par(mfrow = c(1, 1))
qqPlot(lm_feedback_teacher)
```

```
# 3.Create High vs. Low Feedback Literacy Groups
```

```
Data2$F_Lit_Group <- ifelse(Data2$F_LSM_std >= median(Data2$F_LSM_std, na.rm = TRUE), "High", "Low")
```

```
# 4.Compare Feedback Evaluation Between High vs. Low Feedback Literacy Groups (GenAI & Teacher)
```

```
# GenAI (Known)
```

```
t.test(F2A ~ F_Lit_Group, data = Data2)
```

```
# Teacher (Known)
```

```
t.test(F1A ~ F_Lit_Group, data = Data2)
```

```
# 5.Visualization: Feedback Literacy vs. Feedback Evaluation (GenAI & Teacher)
```

```
ggplot(Data2, aes(x = F_LSM_std, y = F2A, color = F_Lit_Group)) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Feedback Literacy vs. GenAI Feedback Evaluation (Known Source)",
```

```

    x = "Standardized Feedback Literacy", y = "Evaluation of GenAI Feedback") +
  theme_minimal()

ggplot(Data2, aes(x = F_LSM_std, y = F1A, color = F_Lit_Group)) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Feedback Literacy vs. Teacher Feedback Evaluation (Known Source)",
    x = "Standardized Feedback Literacy", y = "Evaluation of Teacher Feedback") +
  theme_minimal()

#RQ3
# Standardize AI Literacy (Scale: 1-7)
Data2 <- Data2 %>%
  mutate(AI_LSM_std = scale(`AI-LSM`))

# 1. Correlation Analysis: AI Literacy vs GenAI Feedback Evaluation
cor.test(Data2$AI_LSM_std, Data2$F2B, use = "complete.obs")
cor.test(Data2$AI_LSM_std, Data2$F2A, use = "complete.obs")

# 2.Linear Regression: AI Literacy Predicting GenAI Feedback Evaluation (Known Source)
lm_AI <- lm(F2A ~ AI_LSM_std, data = Data2)
summary(lm_AI)

#Assumption check
model_ai_lit <- lm(F2A ~ AI_LSM_std, data = Data2)

par(mfrow = c(2, 2))
plot(model_ai_lit)

shapiro.test(residuals(model_ai_lit))
ncvTest(model_ai_lit)
par(mfrow = c(1, 1))
qqPlot(model_ai_lit)

# 3.Create High vs. Low AI Literacy Groups Using Median Split
Data2$AI_Lit_Group <- ifelse(Data2$AI_LSM_std >= median(Data2$AI_LSM_std, na.rm =
TRUE), "High", "Low")

# 4.Compare GenAI Feedback Evaluation Between High vs. Low AI Literacy Groups
t.test(F2A ~ AI_Lit_Group, data = Data2)

# 5.Visualization: AI Literacy vs. GenAI Feedback Evaluation (Known Source)
feedback_long <- feedback_long %>%
  mutate(AI_LSM_std = Data2$AI_LSM_std[Participant])

```

```

# Filter for Known condition and AI feedback only
ai_known_data <- feedback_long %>%
  filter(Source == "AI", Awareness == "Known") %>%
  mutate(AI_LSM_std = Data2$AI_LSM_std[Participant])

# Plot
ggplot(ai_known_data, aes(x = AI_LSM_std, y = Rating)) +
  geom_point(alpha = 0.6, color = "#ff7f0e") +
  geom_smooth(method = "lm", se = TRUE, color = "#ff7f0e") +
  labs(
    title = "AI Literacy vs. GenAI Feedback Evaluation (Known Source)",
    x = "Standardized AI Literacy",
    y = "Feedback Rating"
  ) +
  theme_minimal()

# Add AI Literacy Group to ai_known_data
ai_known_data <- ai_known_data %>%
  mutate(AI_Lit_Group = Data2$AI_Lit_Group[Participant])

# Plot with Group Coloring
ggplot(ai_known_data, aes(x = AI_LSM_std, y = Rating, color = AI_Lit_Group)) +
  geom_point(alpha = 0.6) +
  geom_smooth(method = "lm", se = TRUE) +
  labs(
    title = "GenAI Feedback Evaluation by AI Literacy Group (Known Source)",
    x = "Standardized AI Literacy",
    y = "Feedback Rating",
    color = "AI Literacy Group"
  ) +
  scale_color_manual(values = c("High" = "#2ca02c", "Low" = "#d62728")) +
  theme_minimal()

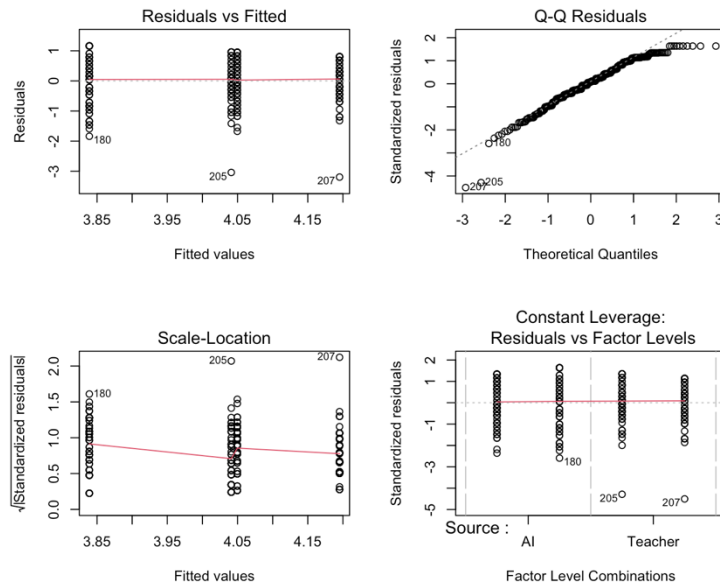
```

## Appendix K

### Assumption Check Plots

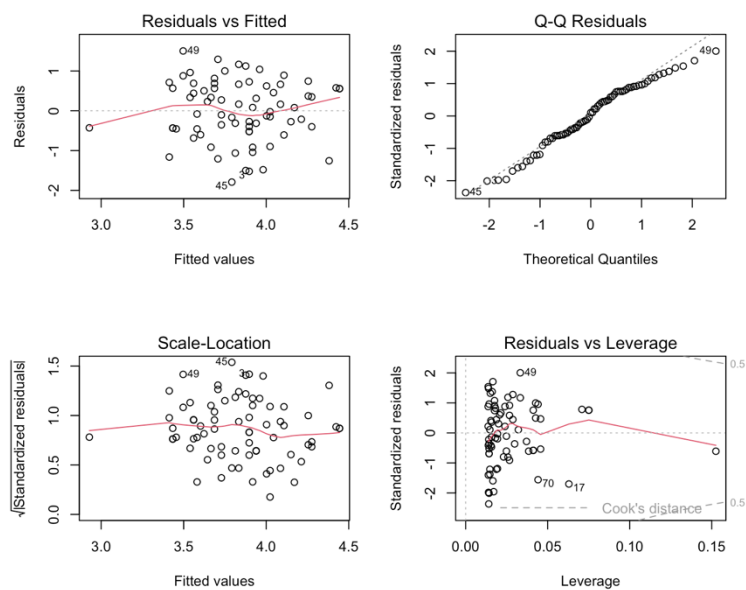
**Figure K.1**

*Q-Q Plot of Residuals for RQ1*

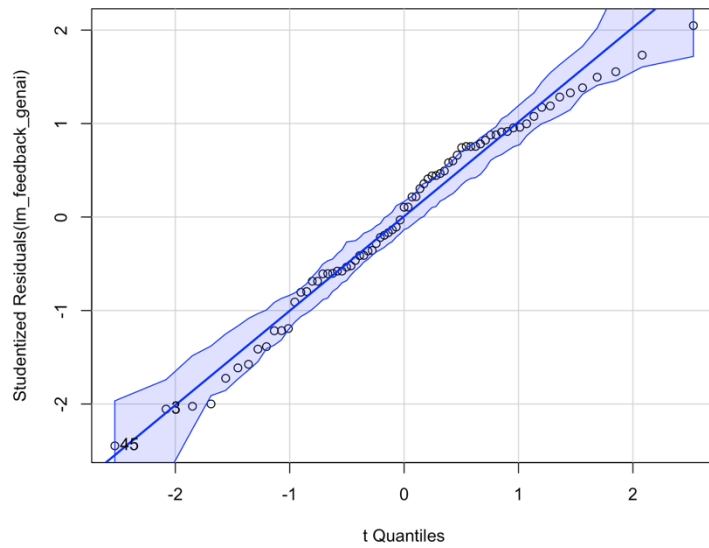
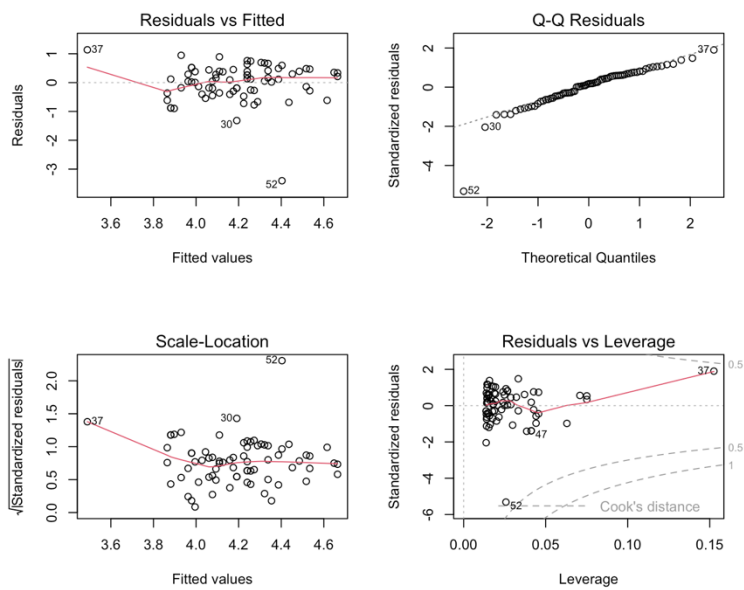


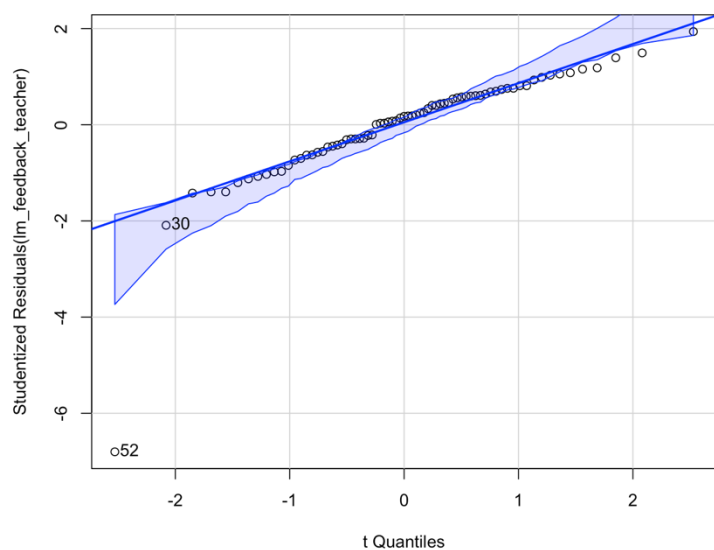
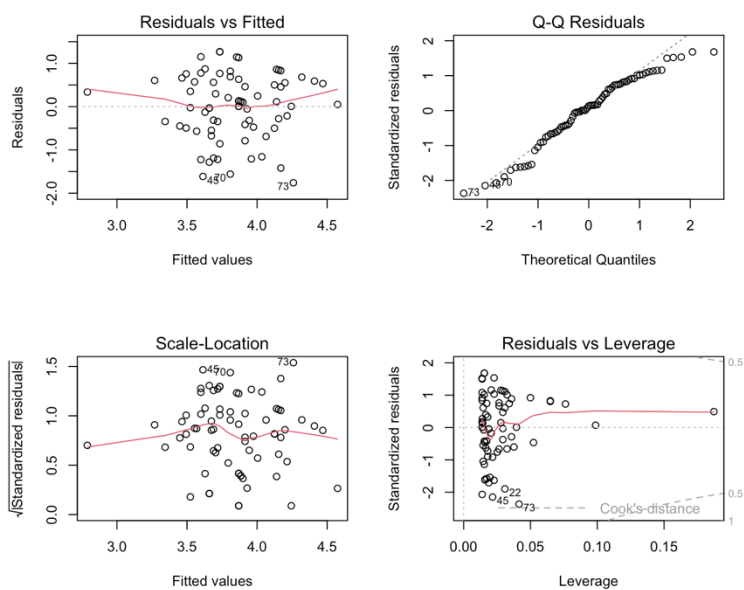
**Figure K.2**

*Q-Q Plot of Residuals for the GenAI Condition for RQ2*





**Figure K.3***Residuals vs Fitted Plot for the GenAI Condition for RQ2***Figure K.4***Q-Q Plot of Residuals for the Teacher Condition for RQ2*

**Figure K.5***Residuals vs Fitted Plot for the Teacher Condition for RQ2***Figure K.6***Q-Q Plot of Residuals for RQ3*

**Figure K.7***Residuals vs Fitted Plot for RQ3*