

HEC MONTRÉAL

Feedback from Tutors to Enhance Second Language Learning:

Exploring the Relationship between Tutor's Feedback

Literacy and AI Acceptance

Par

Farah Hariri

Constantinos K. Coursaris

HEC Montréal

Codirecteur de recherche

Sylvain Sénécal

HEC Montréal

Codirecteur de recherche

Sciences de la gestion

(User Experience)

*Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences en gestion*

(M. Sc.)

Dec 2025

© Farah Hariri, 2025

CERTIFICAT D'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'une évaluation en matière d'éthique de la recherche avec des êtres humains et qu'il satisfait aux exigences de notre politique en cette matière.

Projet # : 2025-6410

Titre du projet de recherche : Feedback to learners to enhance second language learning : Exploring the Relationship between Language Learner's Feedback Literacy and AI Acceptance

Chercheur principal : Sylvain Sénécal, Professeur titulaire, Marketing, HEC Montréal

Cochercheurs : Iman Abdul Razzak; Farah Hariri; Pierre-Majorique Léger; Constantinos K. Coursaris; Marc Fredette; Alexander John Karran; Thaddé Rolon-Mérette; Alejandra Ruiz Segura; Frédérique Bouvier; Luis Carlos Castiblanco; Juan Fernandez Shaw; Andrada Toma; Lucas Dietz; Rayane Benhenni; David Briegne; Salima Tazi; Xavier Côté; Élise Imbeault; Shang Lin Chen

Date d'approbation du projet : 19 mars 2025

Date d'entrée en vigueur du certificat : 19 mars 2025

Date d'échéance du certificat : 01 mars 2026



Maurice Lemelin
Président
CER de HEC Montréal

Signé le 2025-03-19 à 16:10

Abstract

This study examines how tutors conceptualize and deliver feedback in asynchronous second language learning environments, focusing on their acceptance of AI. Using a qualitative single-case design, the study addresses three questions: how tutors define feedback, what are their experiences providing it online, and what are their perceptions of AI's benefits and risks. Data were collected via questionnaires and interviews and analyzed thematically. Findings show tutors view feedback as both cognitive and motivational, guiding learner progress and engagement, but face challenges when delivering asynchronous feedback. Concerns about output accuracy and nuance emerged most prominently. Tutors also mentioned issues related to human touch, privacy, and security, with non-users of AI expressing these concerns more frequently than AI users. Additional themes included anxiety around job displacement, cognitive offloading, and increased surveillance. While AI users tended to emphasize efficiency gains, both groups expressed cautious engagement. The study concludes that developing tutors' feedback literacy alongside careful AI use can improve instructional quality and tutoring, informing discussions on effective feedback and AI integration in language education.

Keywords: Feedback, Learning, Artificial Intelligence, Second Language, Foreign Language, AI-Generated Feedback, Language Learners

Table of Contents

HEC MONTRÉAL.....	1
Abstract.....	2
Table of Contents.....	3
List of Abbreviations and Acronyms.....	4
List of Tables.....	5
Acknowledgements.....	6
Chapter 1.....	7
References.....	16
Chapter 2.....	20
I. Introduction.....	21
II. Literature Review.....	23
III. Methodology.....	32
IV. Results.....	39
V. Discussion.....	51
References.....	59
Chapter 3.....	80
Conclusion.....	80

List of Abbreviations and Acronyms

AI: Artificial Intelligence

ESR: End of Session Report

ESL: English as a Second Language

FSL: French as a Second Language

EFL: English as a Foreign Language

FFL: French as a Foreign Language

TPACK: Technological, Pedagogical, Content, Knowledge

TK: Technological Knowledge

PK: Pedagogical Knowledge

CK: Content Knowledge

TCK: Technological Content Knowledge

TPK: Technological Pedagogical Knowledge

SAS: Statistical Analysis System

List of Tables

Chapter 1

Page 14: Table 1: Contribution to the responsibilities of the research project phases

Chapter 2

Page 34: Table 1. Education and participants, Table 2. Age groups and participants

Page 39: Table 3. Model, Codes, and Participant Quotes

Page 43: Table 4. Perceived Purpose of Feedback Forms Sub-Themes and Representative Participant Quotes

Page 42: Table 5. Perceived Usability of Feedback Forms Sub-Themes, Example Participant Quotes, Codes, and Number of Participants

Page 46: Table 6. Themes, Definitions, Example Quotes and # of Participants

Page 48: Table 7. AI Usage and Themes

Page 49: Table 8. AI Concerns and Technology Knowledge, Table 9. AI Concerns related to AI Usage and Technology Knowledge

Page 50: Table 10. Technology Knowledge Quartiles and Age Groups

Page 51: Table 11. AI Usage in relation to Age

Acknowledgements

I sincerely thank everyone who supported me throughout the development of this thesis. I am particularly grateful to my directors, Dr. Constantinos Cousaris and Dr. Sylvain Sénécal, for their guidance, encouragement, and insightful feedback. I would like to extend my gratitude to the Tech3Lab team, Frédérique Bouvier, Director of UX Researcher, my lab partner, Iman Abdul Razzak, and to the winter cohort that I started with, the old reliables.

Chapter 1

Introduction

1.1 Online Second Language Learning in a Digital World

Acquiring a second language is a multifaceted process that goes beyond learning vocabulary and grammar, requiring learners to develop effective communication skills within new cultural and linguistic settings (Collier, 1995). To support this complex learning process, instructional approaches have evolved significantly, and over the past decade, there has been a marked shift from traditional classroom-based language instruction to online and remote tutoring, driven by technological advances and recent global events such as the COVID-19 pandemic (Moser et al., 2021). Online tutoring can scale to support large numbers of students without requiring proportional increases in physical space or staffing (Parker, 2023). On the other hand, online tutoring can present challenges such as reduced social interaction, technical issues, or the need for self-discipline and digital literacy (Moser et al., 2021). As tutoring increasingly occurs on digital platforms rather than in physical spaces, the interaction between tutor and student becomes a central factor in whether learning is effective.

1.2 The Central Role of Feedback in Second Language Learning

Among the many forms of interaction that support learning, feedback consistently stands out as one of the most effective. Studies show that when feedback is timely, specific, and focused on clear next steps, learners make faster progress, retain information more effectively, and develop greater confidence in their abilities (Wisniewski et al., 2020; OECD, 2025). Whether delivered in person or through digital platforms, feedback helps learners identify misunderstandings, adjust their performance, and strengthen their skills over time (Hattie &

Timperley, 2007). Evidence also indicates that feedback that prompts learner engagement plays a key role in promoting feedback uptake (Lyster & Saito, 2010). Conversely, feedback that lacks clarity, arrives too late, or varies in quality tends to limit its impact and can lead to confusion or discouragement.

1.3 From Online Courses to Asynchronous Tutoring and Spoken Assessment

Digital technologies have transformed educational practices, changing how learners engage with course content, communicate with instructors, and demonstrate their abilities. In the 1960s, the University of Illinois developed the PLATO system, one of the first platforms to provide automated educational feedback, and received instant corrective feedback based on their responses (Bitzer & Skaperdas, 1970). As online learning environments developed, students preferred asynchronous E-learning when synchronous meetings were not possible. This shift contributed to the growth of asynchronous tutoring, where learners receive feedback without needing to meet instructors at specific times (Hrastinski, 2008). However, studies and practitioner reports also point to potential drawbacks, including delays in feedback, inconsistencies in the quality of responses, and differences in how tutors interpret learner work (Jensen et al., 2021).

1.4 AI in Language Education and Feedback

At the same time, Artificial intelligence (AI) is reshaping educational practice. AI has expanded far beyond its early roots in computer science and has become a key technology influencing how digital systems operate and how organizations function. Earlier generations of AI relied heavily on fixed rules and programmed logic, while newer developments in machine learning and deep learning enable systems to learn patterns from extensive datasets and perform tasks that once depended on human expertise (LeCun et al., 2015). A significant milestone in this

evolution was the introduction of the Transformer architecture, which improved the way models handle information and opened the door for advanced language technologies such as GPT (Vaswani et al., 2017).

These innovations have supported the spread of AI across a wide range of areas such as healthcare, finance, government, and customer services, where it is now used to generate predictions, support decision-making, automate routine work, and interact with users (Jordan & Mitchell, 2015). In the context of education, AI tools can come up with topic ideas, practice questions, and assist tutors in writing feedback (Zaimoğlu & Dağtaş, 2025).

With AI appearing more frequently in everyday tools, people increasingly expect these systems to be dependable, understandable, and considerate of user needs (Shneiderman, 2020). As a result, there is growing interest in examining how individuals interpret and respond to AI in both public and private contexts (Long & Magerko, 2020). Research on AI acceptance, building on earlier studies of technology use and trust, explores factors such as how beneficial people believe AI to be, the risks they associate with it, how transparent its processes seem, and the level of trust they place in AI, factors that shape whether these technologies are successfully adopted in real-world environments (Venkatesh et al., 2003; Hoff & Bashir, 2015). These systems are not intended to replace tutors, but instead can support tutors by enhancing clarity, shortening the time needed to draft explanations, and maintaining alignment with learning objectives. When used well, AI can help make feedback more consistent, easier to produce, and more tailored to each learner (Ekizoğlu & Demir, 2025).

1.5. Key Challenges and Research Gap

Despite the growing body of research on feedback in online learning, significant gaps remain. Much of the literature focuses on learners' perceptions and outcomes, with less attention

paid to how tutors define feedback, experience asynchronous feedback delivery, or perceive AI support (Tortella et al., 2025; Chan & Hu, 2023). Among the studies that do exist, many focus on teacher feedback rather than a distinct tutoring role. Moreover, the bulk of studies on asynchronous tutor feedback focus on writing-related feedback, examining how tutors provide feedback on learners' written assignments as opposed to feedback related to spoken assessment (Alnemrat et al., 2025; Mahapatra, 2024; Ekizoğlu & Demir, 2025). These gaps limit our understanding of how feedback practices can be effectively designed and supported in digital learning environments related to spoken assessments.

These gaps have important economic and academic implications. As institutions increasingly rely on asynchronous tutoring and AI-supported feedback to scale language education, there is growing concern that educational technologies may be developed and adopted without sufficient alignment with educators' actual practices (Zawacki-Richter et al., 2019). Addressing these gaps is therefore essential to informing the effective and pedagogically grounded integration of AI in second language learning. Academically, the lack of tutor-focused research limits feedback theory's ability to capture the practical and relational aspects of feedback, which are important for guiding AI-mediated, asynchronous, and spoken assessment (Carless & Boud, 2018).

1.6. Purpose of the Study and Research Questions

The term "feedback" is used in many different ways across research, so it cannot be assumed that tutors share a single, fixed understanding of it. For this reason, it is necessary to explore how tutors personally describe and conceptualize feedback in their own setting. This is especially relevant in second-language tutoring, where the role and intent of feedback may vary compared to other learning contexts. Identifying how tutors themselves understand feedback

forms an essential starting point for this research. Once that shared meaning is articulated, it becomes possible to further explore related questions such as how tutors experience asynchronous feedback delivery and how they perceive AI in relation to their work. RQ1: What is the definition of feedback from a tutor's perspective?

The second objective of this research is to investigate how tutors interact with their internal system to provide feedback to their learners. This entails investigating the procedures tutors use to record comments, scores, or other feedback within the system, along with the obstacles or advantages they experience during this process. Gaining insight into these interactions is important, as the system's design and features can significantly affect the precision and clarity of feedback. Moreover, examining tutors' engagement with the system can shed light on workflow effectiveness, usability concerns, and the influence of technology on tutoring decisions, especially within a second-language learning context. RQ2: What is the user experience of tutors who deliver asynchronous online feedback to learners?

The third objective of this study is to examine tutors' perspectives on AI. This involves exploring their views on the advantages and drawbacks of AI tools both broadly and in the context of their tutoring practice. Understanding these perceptions is especially relevant in second-language instruction, where AI may affect not only administrative activities, such as completing learner reports, but also instructional processes, including providing feedback, supporting learner engagement, and tailoring guidance to individual needs. RQ3: What are the perceptions, perceived risks, and perceived benefits of AI according to tutors?

1.7. Conceptual and Theoretical Foundations

This study builds on Narciss' Model of Feedback (2008), which categorized feedback into 3 main functions: cognitive, motivational, and metacognitive. This study examines how

tutors in an online second language learning environment aligned their definition of feedback. Tutors' responses shed light on the types of feedback they consider most helpful for learners in second language learning contexts.

1.8. Overview of the Study Design

This study included two sources of data: transcripts of individual interviews and questionnaire results. A qualitative case-based research approach was employed, involving interviews with tutors actively engaged in asynchronous online tutoring. It uses theory-informed thematic analysis and questionnaire data as descriptive context. A retrospective think-aloud approach was employed in the secondary part designated to address RQ2, in which participants were shown the feedback form known as the End-of-Session Report (ESR) that they usually filled out and were asked to recall and verbalize how they had completed the tasks. The coding approach was hybrid. Deductively, the interview questions allowed for the creation of high-level categories. Inductively, the responses shaped the creation of themes and codes. After the initial creation of codes, they were organized into broader categories to group together related ideas that might not have been considered initially.

Before the start of the interview, participants were asked to fill out a demographic's questionnaire and a TPACK questionnaire (Schmid et al., 2020). The demographic section of the questionnaire collected information on participants' age, gender, experience with the organization, and educational background. The TPACK questionnaire was designed to measure their self-reported knowledge in 3 main domains: technology, pedagogy, and content knowledge. The quantitative data from the questionnaire were categorized to allow for triangulation with the qualitative data.

This study was limited to participants from a single company. As a result, the findings may not have fully represented the experiences of tutors in other contexts. In addition to that, the small sample size may not fully capture the breadth of experiences and perspectives. Future studies should include participants from diverse institutes and companies, in addition to including more participants to increase the robustness and transferability of the results.

1.9. Contributions and Significance

This study is significant on many levels. Theoretically, it extends the understanding of how tutors define and deliver feedback in an asynchronous second language learning environment, which is under-represented in the literature compared to classroom-based feedback studies. Moreover, this study examines how tutors interact with the systems they use to deliver feedback, highlighting that feedback is not a single action but an ongoing part of their tutoring practice. In addition to that, it contributes to understanding how tutors perceive AI in terms of the advantages and disadvantages in general and within the context of their tutoring practice. The study can be useful to designers, researchers, and institutions looking to implement AI in a way that aligns with tutors' values and pedagogical intent.

1.10. Personal Contribution

Since this thesis was conducted within the Tech3Lab, which included multiple collaborators contributing at different stages and levels, Table 1 outlines my individual intellectual contributions across each aspect of the thesis. In accordance with lab standards, a minimum overall contribution of 50% is expected from the student. Any dimensions where my contributions exceed 50% indicate leadership and ownership of those respective phases.

Table 1: Contribution to the responsibilities of the research project phases

Research Activity	Contribution
Research Questions	Formulating appropriate research questions based on the research partner organization's expectations and needs – 50% Support from the directors and supervisor was provided to determine the research partner's expectations and needs. Support from the directors and supervisor was provided to formulate appropriate research questions.
Study Design	Conceiving and formalizing the study design and interview questions protocol - 60% Members of the Tech3lab conceived the study design.
Questionnaire	Creating the questionnaire on Qualtrics - 90% *Support from the directors and supervisor was provided
Ethics	Requesting ethical approval from the CER - 40%
Recruitment	Recruiting Participants - 50% *Support from the tech3lab and the partner company was provided to recruit participants.
Data Collection	Collecting data - 90% *Support from the team at tech3lab on quantitative collection on Qualtrics was provided
Analyzing	Analyze data - 90% *Support from the directors was provided to guide and revise the results
Writing	Writing introduction, article, and conclusion - 90% *Support from the directors was provided to guide and revise the articles

1.11. Thesis Structure and Positioning of the Article

The thesis is structured into three main chapters, followed by a bibliography and appendix. Chapter 1 introduces the research context, objectives, and significance of the research.

Chapter 2 presents a scientific article, providing a detailed examination of the study's methodology, literature review, data analysis, results, and discussion related to feedback definition, tutor experiences, and AI-related experience. Chapter 3 synthesizes the major findings, demonstrating how the research contributes to existing knowledge, identifies its limitations, and provides suggestions for future research. The bibliography consists of all references, and the appendix includes supplementary materials relevant to the study.

References

- Alnemrat, A., Aldamen, H., Almashour, M., Al-Deaibes, M., & AlSharefeen, R. (2025). AI vs. teacher feedback on EFL argumentative writing: A quantitative study. *Frontiers in Education, 10*, 1614673. <https://doi.org/10.3389/feduc.2025.1614673>
- Bitzer, D. L., & Skaperdas, D. (1970). A large scale computer-based educational system for the seventies. In *Pace Congress® Proceedings 1970 (7th), Technology Today and Tomorrow*. University of Illinois, Urbana.
- 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>
- 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), 43. <https://doi.org/10.1186/s41239-023-00411-8>
- Collier, V. P. (1995). *Acquiring a second language for school: A conceptual model*. National Clearinghouse for Bilingual Education.
- Ekizoğlu, M., & Demir, A. N. (2025). The role of AI-assisted writing feedback in developing secondary students’ writing skills. *Discover Education, 4*, 454. <https://doi.org/10.1007/s44217-025-00919-3>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hoff, K., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 57*(3), 407–434. <https://doi.org/10.1177/0018720814547570>

- Hrastinski, S. (2008). Asynchronous and synchronous e-learning. *EDUCAUSE Quarterly*, 31(4), 51–55.
- Jensen, L & Bearman, M & Boud, D. (2021). Understanding feedback in online learning – A critical review and metaphor analysis. *Computers & Education*. 173. 104271. 10.1016/j.compedu.2021.104271.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- Parker, D. (2023). *Scaling tutoring: Five key factors to benefit more students*. Inter-American Development Bank. <https://publications.iadb.org/en/scaling-tutoring-five-key-factors-benefit-more-students>
- Schmid, M., Brianza, E., & Petko, D. (2020). Developing a short assessment instrument for Technological Pedagogical Content Knowledge (TPACK.xs) and comparing the factor structure of an integrative and a transformative model. *Computers & Education*, 157, 103967. <https://doi.org/10.1016/j.compedu.2020.103967>
- Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human–Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32(2), 265–302. <https://doi.org/10.1017/S02722263109990520>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>

- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(1), 9.
<https://doi.org/10.1186/s40561-024-00295-9>
- Moser, K. M., Wei, T., & Brenner, D. (2021). Remote teaching during COVID-19: Implications from a national survey of language educators. *System*, 97, 102431.
<https://doi.org/10.1016/j.system.2020.102431>
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 125–144). Lawrence Erlbaum Associates/Taylor & Francis.
https://www.researchgate.net/publication/233833269_Feedback_Strategies_for_Interactive_Learning_Tasks
- Long, D., & Magerko, B. (2020). *What is AI literacy? Competencies and design considerations*. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery.
<https://doi.org/10.1145/3313831.3376727>
- OECD (2025), *Unlocking High-Quality Teaching*, OECD Publishing, Paris,
<https://doi.org/10.1787/f5b82176-en>
- Tortella, F., Palese, A., Turolla, A., Castellini, G., Pillastrini, P., Landuzzi, M. G., & Rossetti, G. (2025). Knowledge and use, perceptions of benefits and limitations of artificial intelligence chatbots among Italian physiotherapy students: A cross-sectional national study. *BMC Medical Education*, 25, 572. <https://doi.org/10.1186/s12909-025-07176-w>

- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, 3087.
<https://doi.org/10.3389/fpsyg.2019.03087>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*, 5998–6008.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly, 27*(3), 425–478.
<https://doi.org/10.2307/30036540>
- Zaimoğlu, S., & Dağtaş, A. (2025). Teacher Cognition and Practices in Using Generative AI Tools to Support Student Engagement in EFL Higher-Education Contexts. *Behavioral Sciences, 15*(9), 1202. <https://doi.org/10.3390/bs15091202>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, 1–27. <https://doi.org/10.1186/s41239-019-0171-0>

Chapter 2

Feedback from Tutors to Enhance Second Language Learning: Exploring the Relationship between Tutor's Feedback Literacy and AI Acceptance

Abstract

This study explores how tutors conceptualize feedback and experience its delivery within asynchronous second language learning environments, with particular attention to their perceptions of artificial intelligence (AI) technologies. Adopting a qualitative single-case study design, the research investigates three key questions: how tutors define feedback, what user experiences they encounter when providing online feedback, and how they perceive the potential benefits and risks of AI integration. Data were collected through questionnaires and in-depth qualitative analysis using a theory-informed thematic approach. Findings indicate that tutors primarily view feedback as both a cognitive and motivational tool, emphasizing its role in guiding learner progress and sustaining engagement. They reported challenges related to delivering asynchronous feedback, such as a perceived lack of learner engagement. Regarding AI, tutors expressed cautious optimism, acknowledging its potential to enhance efficiency and consistency while voicing concerns over a lack of human touch. The study concludes that fostering tutors' feedback literacy alongside thoughtful AI adoption can enhance both instructional quality and tutoring. These insights contribute to the evolving discussion on effective feedback practices and the integration of AI in language education.

Keywords: Feedback, Learning, Artificial Intelligence, Second Language, Foreign Language, AI-Generated Feedback, Language Learners

I. Introduction

Feedback is widely acknowledged as a core component of effective learning, guiding learners' understanding, motivating improvement, and shaping educational outcomes (Lilly et al., 2010). In online learning environments, particularly asynchronous contexts, feedback delivery is increasingly mediated through digital platforms, which introduces unique challenges for tutors. Unlike traditional face-to-face interactions, asynchronous feedback often lacks immediate dialogue, non-verbal cues, and real-time clarification, potentially affecting both the quality of feedback and tutors' experiences (Mandouit & Hattie, 2023).

Understanding tutors' perspectives on feedback is essential because they are responsible for guiding learners' progress while navigating technological, pedagogical, and time-related constraints in online teaching (Bell, 2019). Tutors' experiences can influence the clarity, usefulness, and timeliness of feedback, which in turn affects learner engagement and satisfaction. Additionally, the integration of artificial intelligence (AI) into feedback processes adds further complexity, raising questions about tutors' trust in AI tools, perceived risks, and potential benefits in supporting pedagogical practices (Wang & Li, 2024).

Despite the growing body of research on feedback in online learning, significant gaps remain. Much of the literature focuses on learners' perceptions and outcomes, with less attention paid to how tutors define feedback, experience asynchronous feedback delivery, or perceive AI support (Tortella et al., 2025; Chan & Hu, 2023). Among the studies that do exist, many focus on teacher feedback rather than a distinct tutoring role. According to Walker (2018), tutoring is understood as a supportive, learner-focused role distinct from formal teaching, emphasizing guidance, mentoring, and individualized assistance rather than curriculum delivery.

In addition to that, the bulk of studies on asynchronous tutor feedback focus on writing-related feedback, examining how tutors provide feedback on learners' written assignments as opposed to feedback related to spoken assessment. (Alnemrat et al., 2025; Mahapatra, 2024) These gaps limit our understanding of how feedback practices can be effectively designed and supported in digital learning environments.

To address these gaps, this study explores three interrelated research questions:

1. What is the definition of feedback from a tutor's perspective?
2. What is the user experience of tutors who deliver asynchronous online feedback to learners?
3. What are the perceptions, perceived risks, and perceived benefits of AI according to tutors?

A qualitative case-based research approach was employed, involving interviews with tutors actively engaged in asynchronous online tutoring. Second-language tutors were interviewed to examine how they define feedback, how they used the partner company platform to deliver feedback which was required post-session, and how they perceived AI, including their perspectives, experiences, and attitudes toward AI integration.

This study contributes to the literature in three main ways. First, it clarifies how tutors conceptualize feedback in asynchronous online contexts, whereby it was found that tutors emphasize motivation as an important part of providing feedback. This aligns with Narciss' (2008) model of feedback while extending the literature by considering that in language learning environments, motivational support might play a larger role than metacognitive feedback.

Second, this research highlights the experiential challenges tutors face with their company platform when providing feedback. The findings show that tutors perceive the feedback

forms to primarily serve administrative requirements and tutor-related functions over learner development. This extends the work of Sadler (2010) by identifying a third purpose of feedback forms to include tutor-related functions. In addition, tutors reported that the Numeric Rating Scales within the feedback forms made it difficult to accurately assess learners' skill levels. These insights offer concrete guidance for designing more supportive and pedagogically aligned feedback systems.

Third, the research provides empirical evidence on tutors' perceptions of AI whereby it was found that AI usage or lack thereof can give rise to different concerns, suggesting that concerns are shaped by exposure to the technology. Tutors with higher or lower self-assessed technological knowledge also mentioned different AI-related concerns, indicating that perceptions of AI can also be tied to one's sense of digital competence, not just to the technology exposure itself. Collectively, these findings advance research on online feedback by centering tutors' perspectives, addressing gaps in the literature, and offering practical implications for enhancing both tutor and learner experiences.

II. Literature Review

Feedback Literacy in Language Learning

Feedback is information given to learners about their performance with the goal of guiding improvement and supporting learning (Shute, 2008). It goes beyond simply indicating correctness, helping learners reflect, adjust strategies, and develop self-regulation. Sadler (1989) emphasizes that effective feedback requires learners to understand the desired standard, recognize gaps in their performance, and receive guidance on how to close those gaps. While feedback once mostly implied teacher's comments on student work, the rise of digital learning has meant that it is now also used to describe many forms of automatically generated

information. For example, online quizzes with pre-set commentaries are often considered feedback (e.g., Forster et al., 2018; Maier et al., 2016). Overall, feedback is both corrective and developmental, supporting learning through guidance, reflection, and engagement.

Research by Jenson et al. (2023) has identified 6 metaphors for feedback. These metaphors are (1) feedback is a treatment, (2) feedback is a costly commodity, (3) feedback is coaching, (4) feedback is a command, (5) feedback is a dialogue, and (6) feedback is a learner tool. Of these, the coaching metaphor is particularly relevant to this study, as it emphasizes the motivational and emotional role of feedback in online learning contexts, where in-person interaction is often limited. This perspective emphasizes the emotionally charged nature of tutor–learner interactions and the role of feedback in sustaining motivation. However, this framing also raises concerns that emphasizing motivation might overshadow the pedagogical intent of feedback (Jensen et al., 2021).

Another relevant approach in providing feedback is known feedback sandwich, where a person first highlights positive aspects, then addresses areas for improvement, and concludes with positive reinforcement, leveraging the motivating effect of praise to make constructive criticism easier to accept (Dohrenwend, 2002). The goal is to make feedback more acceptable and motivating, helping learners respond constructively to suggestions. While it is often applied in classroom and workplace settings, the model highlights how the presentation of feedback can influence how it is interpreted and acted upon. Research by Prochazka et al., (2020) demonstrates that students perform better in the next task if they receive sandwich feedback than if they receive no feedback or corrective feedback only. However, some scholars caution that by sandwiching criticism between praise, the negative feedback may lose its impact. Recipients

might focus more on the positive comments and overlook the areas needing improvement, leading to missed opportunities for growth (Bressler & Von Bergen, 2014).

Over the years, scholars have developed multiple models to explain how feedback functions and how it can be optimized in educational and performance contexts. Feedback models are theoretical frameworks that explain how feedback works, why it is effective, and the processes through which it supports learning or performance improvement (Hattie & Timperley, 2007; Sadler, 1989). These models provide guidance for designing, structuring, and delivering feedback in ways that promote understanding, engagement, and skill development. Rather than focusing on the specific form feedback takes, models describe the underlying mechanisms, principles, and functions that make feedback effective. The following section reviews several influential feedback models that illustrate different approaches to guiding learners toward improvement.

One model of feedback particularly relevant to this study is Narciss' (2008) Model of Feedback since this study uses it as a theoretical foundation. In order to understand feedback beyond simply telling learners what was right or wrong, Narciss (2008) conducted research and came up with a model that categorized feedback into 3 main functions. Those functions are cognitive, motivational, and metacognitive. The cognitive function of feedback provides task-related information that helps learners close the gap between actual and desired performance by clarifying errors, correct responses, and effective strategies. The motivational function of feedback concerns its role in influencing learners' effort, persistence, and attitudes toward learning by supporting self-efficacy and sustaining engagement. The third function from Narciss's model is the metacognitive function. It is described as guiding learners in monitoring, evaluating, and regulating their own learning processes, thereby fostering self-regulated learning.

Beyond understanding what function feedback serves, it is also relevant to understand how educators perceive what purpose it aims to fulfill, as these two are not always in conjunction with one another. According to Sadler (2010), feedback fulfills two interconnected yet often conflicting roles: it supports student learning and development, while also serving an administrative or evaluative function. This dual purpose creates a tension between promoting growth and maintaining accountability within educational assessment systems. The developmental aspect of feedback centers on fostering dialogue, critical reflection, and learner independence, framing feedback as a means for continual growth. Conversely, its administrative role connects feedback to institutional processes such as grading, evaluation, and accountability, often serving as proof of instructional quality or compliance. Sadler (2010) notes that these two purposes can create friction, as an emphasis on evaluation may weaken the formative value of feedback. When feedback is primarily used to defend grades or fulfill institutional expectations, students are more likely to interpret it as judgmental rather than developmental, which can limit their engagement with its learning potential.

Feedback Artefacts and Practical Challenges

While feedback models describe the mechanisms and rationale behind feedback, and feedback purpose describes the organizational challenges involved with ensuring that feedback serves learner development, feedback types refer to the observable forms, formats, or categories of feedback that learners receive in practice (Shute, 2008). Research in educational technology highlights that feedback can take multiple forms, each serving different instructional purposes (Shute, 2008). This section examines the main categories of feedback, emphasizing how they are applied to support learning, skill development, and self-regulation. For clarity, types of feedback can be organized across several dimensions: timing, focus, function, and modality. Engaging

with this literature is important for the study because it clarifies the conceptual and practical foundations underlying the feedback form used in the research.

Feedback can be delivered either immediately or after a delay. Immediate feedback is provided directly following learner responses, allowing errors to be corrected in real time and reinforcing correct language patterns (Shute, 2008). Delayed feedback, on the other hand, is provided to a learner after a time interval following their response or performance, rather than immediately. (Metcalf et al., 2009) End-of-session summaries in online tutoring exemplify this approach.

According to Lyster and Ranta (1997), feedback can also be categorized by focus. Corrective feedback can be categorized into three types: recasts, in which a teacher reformulates a student's error without explicitly indicating it; explicit correction, which provides the correct form while clearly signaling the error; and prompts, which give hints or cues that encourage the student to self-correct.

Feedback can serve different functions within learning and assessment processes. The literature commonly distinguishes between formative feedback, which functions to guide improvement during learning, and summative feedback, which functions to evaluate performance after completion (Sadler, 1989; Hattie & Timperley, 2007). Formative assessment does not take the form of a particular instrument or task (Moss & Brookhart, 2019), but is defined by its purpose, which is to help form, or shape, a student's learning during the learning process (Shepard, 2000). It can even bolster students' abilities to take ownership of their learning when they understand that the goal is to improve learning, not apply final marks (Trumbull and Lash, 2013). Summative or evaluative feedback, commonly associated with grades or assessments, serves to measure performance but may not directly enhance learning (Black & Wiliam, 1998).

Feedback may also differ in its mode of delivery. Written feedback, including text-based comments or corrections, allows learners to review guidance at their own pace. Oral or audio feedback supports pronunciation, intonation, and oral communication skills, while visual feedback uses colors, highlights, or interactive charts to signal errors and track progress (Shute, 2008).

To enrich the examination of feedback types, this section introduces key usability principles, which provide a foundation for evaluating how effectively feedback mechanisms support user needs in the case presented in this study. This study draws on ISO 9241-11 (2018), which defines usability through three key dimensions: effectiveness, efficiency, and satisfaction. In usability research, effectiveness is defined as the accuracy and completeness with which users achieve specified goals in particular contexts of use. Satisfaction in usability is defined as the degree to which a system provides comfort and acceptability to users in its context of use. Efficiency refers to the resources expended in relation to the accuracy and completeness with which users achieve goals (ISO 9241-11, 2018).

This study adopted a think-aloud protocol to gain insight into users' thought processes in order to assess system usability, as it helps reveal errors, inefficiencies, and areas of confusion which is why. This protocol is commonly employed in usability research to capture users' reasoning and decision-making while interacting with an interface (Ericsson & Simon, 1993). Through verbalizing their thoughts, participants offer information about a system's effectiveness, efficiency, and satisfaction (ISO 9241-11, 2018). Both concurrent and retrospective think-aloud approaches have been utilized in previous studies: the former records cognition in real time during task performance, while the latter depends on participants' recollection of tasks already

completed (Haak et al., 2003). Examining these methods establishes the conceptual basis for the methodological choices made in this study.

AI Perceptions, Knowledge, and Acceptance

Particularly relevant to asynchronous and formative feedback is the development of AI tools. Recent research shows that AI-generated feedback is increasingly used as a formative tool in ESL/EFL settings, with generally positive effects on writing skill development (Mahapatra, 2024). Another study indicated that AI-generated feedback can lead to significant improvements in language learners' writing skills, comparable to traditional teacher feedback (Alnemrat et al, 2025). When interacting with teacher-generated feedback, learners often benefit from interpersonal trust, contextual knowledge, and nuanced scaffolding that is aligned with classroom dynamics. However, such feedback may be delayed due to time constraints and can sometimes be inconsistent in tone or focus (Zhang & Hyland, 2018). By contrast, AI-generated feedback offers immediacy and consistency, and can be tailored through prompt engineering to target specific rhetorical or genre-related issues (Guo et al., 2024). Yet, studies show that learners may interact with AI feedback more passively, often accepting suggestions without critical evaluation or reflection (Yoon et al., 2023). This passive uptake raises concerns about surface-level revision, overreliance, and reduced metacognitive engagement. However, the limitations of AI feedback are also widely acknowledged. Yoon et al., (2023) caution that while ChatGPT can generate plausible feedback on coherence and logic, it may also “hallucinate” critiques, offer generic advice, or miss contextually nuanced issues.

The growing integration of AI into education has sparked both optimism and concern among teachers. Studies suggest that educators' perceptions significantly shape how these technologies are adopted and implemented in classrooms (Alshahrani & Mostafa, 2025).

Generally, tutors view AI as a tool that can improve teaching efficiency and learning outcomes, although they remain wary of its broader effects on pedagogy, ethics, and the teaching profession.

Educators frequently highlight AI's usefulness in reducing administrative burdens, facilitating individualized learning, and improving assessment processes (Zhai & Nehm, 2023). Alshahrani and Mostafa (2025) found that university instructors perceive AI systems as beneficial for lightening workload and enhancing feedback quality. While acknowledging AI's advantages, teachers also expressed apprehension about its educational and ethical consequences. Ethical concerns, including privacy, algorithmic bias, and transparency, regularly appear in the literature as key barriers to acceptance (Holmes et al., 2022). These concerns are amplified by a lack of training and limited understanding of how AI-generated decisions are formed (Zhai & Nehm, 2023).

Overall, the body of research reflects a sense of measured optimism among educators. Teachers recognize the potential of AI to enhance productivity and personalization in learning but remain cautious about the ethical risks and its influence on the human aspects of education. Effective adoption of AI, according to scholars, relies on comprehensive professional training, transparent governance, and institutional strategies that preserve the human-centered values of teaching (Holmes et al., 2022; Alshahrani & Mostafa, 2025).

It is particularly important to understand what contributes to the perceptions of AI since it can play a role in whether someone adopts it or not. From a psychological standpoint, increased exposure to knowledge of a technology typically enhances users' comfort and confidence with that tool, though this relationship has important nuances. Recent scholarship indicates that individuals with stronger AI literacy tend to report higher confidence levels, more positive

perceptions of AI's utility, and more favorable overall attitudes toward these systems (Bewersdorff et al., 2025), though this correlation does not follow a simple linear trajectory. Pan et al. (2025) identify a phenomenon they call "informed skepticism," where users with deeper AI knowledge become more demanding and critical in their trust assessments, particularly when they recognize concerns such as algorithmic bias, opaque decision-making processes, or unreliable outputs. This dual pattern is consistent with established technology acceptance frameworks, which suggest that competence can simultaneously increase perceptions of a technology's value while decreasing user anxiety (Compeau & Higgins, 1995), while also supporting the concept of knowledgeable users functioning as "cautious critics" who pair their expertise with careful scrutiny (Bewersdorff et al., 2025).

Recent research delved deeper into how AI knowledge and confidence shaped participants' AI risk-benefit perception (Said et al., 2023). People with higher AI knowledge exhibited "risk blindness", meaning they underestimated the risks associated with AI applications. This suggests that more knowledge does not necessarily lead to a more accurate assessment of risks and benefits. In contrast, higher confidence in one's AI knowledge led to a larger difference between risk and benefit ratings, biasing people towards either overestimating the risks or overestimating the benefits.

Particularly relevant to the measurement of pedagogy and technology knowledge is a self-assessment questionnaire known as the Technological Pedagogical Content Knowledge Framework (TPACK). The TPACK framework, introduced by Mishra and Koehler (2006), conceptualizes the knowledge teachers need to integrate technology effectively into their teaching. This instrument has been adapted and employed extensively in teacher education and technology integration research across different cultural and educational contexts (Chai et al.,

2013). It is important to note that the TPACK framework is not focused on any particular technology, it is technology agnostic (Mishra et al., 2023), making it an amenable framework for considering how AI can be used to support teaching and learning.

III. Methodology

1. Research Method and Context

This qualitative single-case study aimed to explore the tutors' perceptions, experiences, and potential pain points while utilizing a language learning platform with implications for improving feedback delivery to learners. Yin (2018) suggests that a single-case study enables researchers to thoroughly explore a present-day phenomenon in its natural context, particularly when it is difficult to separate the phenomenon from its surrounding conditions. In such situations, understanding the case requires examining how organizational or social factors interact with the phenomenon under study. This study uses theory-informed thematic analysis and questionnaire data as descriptive context. It suggests that understanding tutor experiences could inform the development of better online formative feedback methods, while improving tutors' work experiences, and strategically preparing for AI's role, ultimately boosting learner outcomes. A qualitative approach was selected to explore participants' perceptions in depth, which aligns with the study's aim of understanding tutor experiences.

The research study was carried out with a North American company that provides online tutoring sessions for adult learners to help prepare them for official language proficiency exams. The tutoring company employed a module-based approach in which learners first engaged with prepared online content independently, and were then scheduled a certain number of tutoring sessions per week (either one-on-one with the tutor or with a group of three other learners) to

practice and refine their understanding of the material and the language of study. The modules and tutoring sessions cater to learners at the beginner, intermediate, and advanced levels.

It is also important to note that after every tutoring session, an end-of-session report (ESR), also known as the feedback form, was filled out by the tutor to assess learners on their linguistic knowledge and their performance (*see Appendix A for the partner company's ESR.*) The linguistic elements assessed were labeled as follows: Conjugation, Syntax, Vocabulary, Spelling, Pronunciation, and Comprehension. There was an open text field for each of these elements for the tutor to fill in their comments and a subsequent section known as learner's performance, where tutors could rate the learner on a 1 to 5 scale, 1 being weak and 5 being strong, for each linguistic element. There was also an open-ended text field where tutors could input the topic of the session.

2. Participants

The participants were recruited with the assistance of the partner company, which sent out a form to their employees (tutors) informing them of the voluntary study. Once this list was received, the author reached out to the interested participants. All participants were informed of the purpose of the study and their right to withdraw at any time. All participants signed a consent form agreeing to an audio recording and were assured of their anonymity. To maintain confidentiality, participants were assigned a participant number and identifying information was removed from the transcripts. HEC's research ethics board provided approval, under the certificate number #2025-6410.

As a result, 15 second language tutors from the partner company participated in this study. Tutors of all backgrounds, years of experience at the organization, educational levels, and ages were invited to partake in this study. Thirteen participants were French language tutors,

while 2 were English language tutors. Twelve participants had a career in teaching, while 3 came from other fields of study in Science, Technology, Engineering, and Math (STEM). Participants' educational backgrounds were spread throughout all levels. Table 1 below summarizes the education level and number of participants attributed to each.

Table 1. Education and participants

Educational backgrounds	n
Post Graduate Degree	2
Graduate Degree	6
Undergraduate Degree	3
Undergraduate Certificate	2
College Degree	2

Participant ages ranged from 29 to 77. Five participants were in their 20s-30s, 4 participants were in their 50s, 5 participants were in their 60s-70s. Table 2 below summarizes participant age groups. These age groups were based on Arnett's (2000) concept of emerging adulthood.

Table 2. Age groups and participants

Age Range	n
29-37	6
51-58	4
61-77	5

3. Measures

To ensure the rigor of this study, several measures were implemented. Credibility was enhanced by reviewing coding and findings with 4 colleagues to verify interpretations. Dependability was supported through a comprehensive record of all coding decisions.

Confirmability was achieved by keeping reflective notes and analytic memos to ensure that findings were firmly based on the data. Transferability was addressed by providing demographic information about participants and their background, allowing readers to determine whether the findings could apply in similar situations.

This study included two sources of data: transcripts of individual interviews and questionnaire results. The study adopted a semi-structured interview approach to address the research questions. Interviews are particularly useful to collect information, which fails to be observed directly (*see Appendix B for the complete interview guide*). As such, it is a widely used instrument to elicit teachers' beliefs, since beliefs reside in teachers' minds and are unseen thought processes (Pajares, 1992). The semi-structured interviews were conducted on Microsoft Teams, each lasting 45-50 minutes, and were audio-recorded with consent.

A retrospective think-aloud approach was employed in the secondary part designated to address RQ2, in which participants were shown the feedback form or ESR and were asked to recall and verbalize how they had previously filled it out. Although the ESR could have been filled out just before the interview was conducted, the verbalizations reflect retrospective recall rather than real-time task execution, distinguishing this method from classical concurrent think-aloud protocols and providing insight into users' reasoning and decision-making during past interactions (Haak et al., 2003; Ericsson & Simon, 1993). This approach allowed the research to capture participant reasoning and potential usability issues without requiring task performance. To address RQ3, in which participants were asked about their thoughts related to AI, information was collected on their usage of AI, the context of their use and their concerns regarding AI.

Before the start of the interview, participants were asked to fill out a demographic's questionnaire and a TPACK questionnaire (Schmid et al., 2020), which is widely used to assess teachers' technological, pedagogical, and content knowledge. This TPACK version has been empirically tested and validated, making it appropriate for measuring Technological Pedagogical Content Knowledge in this study (Schmid et al., 2020). The TPACK questionnaire measured participants' self-reported knowledge in 3 main domains, technology, pedagogy, and content knowledge (Mishra & Koehler, 2006, *see Appendix C for the measurement scale used for the TPACK questionnaire*). The questionnaire consists of self-report items rated on a Likert scale (e.g., 1 = strongly disagree, 7 = strongly agree). For example, an item from the pedagogical knowledge sub-scale is: *"I can adapt my teaching based upon what students currently understand or do not understand."*

The demographic section of the questionnaire collected information on participants' age, gender, years of work experience with the organization, and educational background. Differences in age appear to influence how people interpret, trust, and respond to feedback provided by AI systems. Research indicates that younger users tend to demonstrate higher competence with AI tools and greater trust in AI-generated feedback, while older users often adopt more cautious or critical stances. For example, a recent survey found that participants under age 49 had a higher median self-rated AI competence compared to those aged 50 and over (Lāma & Lastovska's, 2025). The age groupings used in this study, young adults (20s–30s), midlife adults (40s–50s), and older adults (60s–70s), are based on Arnett's (2000) concept of emerging adulthood. According to Arnett (2000), emerging adulthood is a developmental period between adolescence and full adult independence, generally encompassing the late teens through the twenties. This perspective emphasizes that becoming an adult is a gradual process that often continues into the

late twenties and early thirties, supporting the designation of the young adults (20s–30s) category. Extending this view, the framework also informs the classification of midlife (40s–50s) and older adults (60s–70s), reflecting shifts in life responsibilities, social roles, and psychosocial priorities across adulthood.

4. Data Analysis

Thematic analysis was used to identify patterns and key themes from the interview data. According to Braun and Clarke (2006), thematic analysis offers an accessible and theoretically flexible approach to analyzing qualitative data. To prepare the data for analysis, transcripts were extracted from teams and uploaded to Microsoft Word, they were manually cleaned and anonymized and converted to txt format to be uploaded on NVivo. The transcripts were then coded on NVivo. Coding was applied to meaningful verbatim, typically between one sentence and one short paragraph, which represented a single idea expressed by the participant.

The data were analyzed using Braun and Clarke's (2006) six-phase framework for thematic analysis. Initially, transcripts were read multiple times to become familiar with the content and generate preliminary ideas. Significant elements were then systematically coded, reflecting both surface-level content and deeper meanings. These codes were organized to identify potential themes, which were subsequently reviewed and refined to ensure clarity and distinctiveness. Each theme was then clearly defined and named, and the final analysis was produced by weaving the themes together with supporting data extracts, linking the findings to the research questions. This iterative approach allowed for a thorough and transparent interpretation of the qualitative data.

The coding approach was hybrid. Deductively, the interview questions allowed for the creation of high-level categories. Inductively, the responses shaped the creation of themes and

codes. After the initial creation of codes, they were organized into broader categories to group together related ideas that might not have been considered initially. Categories were then refined into overarching themes that captured key dimensions of participants' perspectives.

The quantitative data from the questionnaire were categorized to allow for triangulation with the qualitative data. In addition, participants were categorized into three age groups to facilitate analysis of age-related patterns and to explore potential differences in perceptions of AI-generated feedback. The ages of sample interviewed ranged from 29 to 77 years, and participants were grouped based on clusters of similar ages informed by research by Arnett (2000). The first group included younger adults aged 29–37 (6 participants), the second group comprised middle-aged adults aged 51–58 (4 participants), and the third group consisted of older adults aged 61–77 (5 participants). See Appendix D for all participants' ages. This grouping approach prioritized meaningful age clusters, allowing for clearer interpretation of differences between age groups while reflecting the distribution of participants in the sample.

The data from the TPACK questionnaire were analyzed using SAS. The results were then merged into clusters for each dimension. The mean of all technology-related scales (TK, TCK, TPK) were extracted, then quartiles were created using the median-of-halves using Tukey method in order to find participants that were on the high/medium/low quartiles. Attributes were then created for each participant on NVivo by assigning a quartile to their profile. This allowed for cross-referencing themes and codes with their scores of self-reported knowledge.

The thematic analysis for the research question “What is the definition of feedback from a tutor’s perspective?”, was guided by Narciss’s (2008) model, which categorized feedback functions into cognitive, motivational, and metacognitive domains. Because this was a qualitative study and no tasks took place, the 2nd research question “What is the user experience

of tutors who deliver asynchronous online feedback to learners?”, utilized the ISO usability framework which breaks down usability attributes into 3 components: effectiveness, efficiency, and satisfaction (ISO 9241-11, 2018). The attributes were coded in the following manner: (1) Effectiveness: when a user mentioned struggles with accuracy. (2) Satisfaction: when a user mentions frustration or comfort. (3) Efficiency: when a user estimates effort/time to fill out the ESR. The results of the time to complete the ESR were coded into four categories (very short, short, medium, long) to identify patterns of efficiency across participants. Very short was considered 2 minutes, short was considered to be 5 minutes, medium was considered between 6 and 10 minutes, long was considered more than 10 minutes but less than 15.

Finally, to analyze the results for the research question “What are tutors’ perceptions, experiences, concerns and their perceived benefits of AI?”, participant concerns were mapped to whether they belong to participants who use or don’t use AI for tutoring. AI Usage was also mapped to the participants’ age group. Finally, technological knowledge, which was collected from the TPACK questionnaire, was mapped to AI usage, AI Concerns, and Participant Age Groups.

IV. Results

1. Tutors’ definition of feedback in the context of asynchronous second language tutoring

In addressing the research question “How do tutors define feedback?”, three main codes emerged from the analysis that explained how participants defined feedback. These codes were: (1) *Cognitive Function*, (2) *Motivational Function*, and (3) *Metacognitive Function* (Narciss, 2008). The following sections present each code in detail, supported by participants’ verbatim. Table 3 shows the coding scheme for the definition of feedback using Narciss’ Model of Feedback (2008).

Table 3. Model, Codes, and Participant Quotes

Model	Code	Example
Narciss' Model of Feedback	A. Cognitive Function	"Feedback is information that you give to another individual about something that they have said or practiced to let the person know how they're doing..." (P08)
	B. Motivational Function	"It can also act as encouragement... to encourage learners to continue and to keep doing what they're doing" (P15).
	C. Metacognitive Function	"I can ask my learner for feedback on his oral evaluation, especially when he failed his oral evaluation, I can ask him the following session: Give me your feedback about your experience for example" (P03)

The findings for RQ1 suggested that tutors predominantly aligned their definition of feedback with cognitive and motivational functions. In contrast, feedback aimed at fostering learner self-reflection (a key aspect of the metacognitive function) appeared to be less emphasized in the context of second language tutoring.

All participants identified the act of informing learners about errors as a core component of feedback which falls under the cognitive domain. However, several participants expanded their definitions to include additional functions. Notably, seven participants emphasized the importance of delivering feedback in a supportive and encouraging tone, highlighting its motivational role. Furthermore, four participants, three of whom also mentioned motivational aspects, introduced a metacognitive dimension by suggesting that feedback should prompt learners to reflect on their learning strategies.

As the conversation carried on, it became apparent that the motivational function played an important role in the long-term goal of language acquisition. Following up with tutors with

the question “What is your approach to giving feedback?” yielded more responses related to motivation; ten out of fifteen participants described using positive framing and the “feedback sandwich” technique. Only 4 participants implicitly referenced the metacognitive function’s role in tutoring. When participants were asked about their approach to giving feedback, only one additional participant implicitly referenced the metacognitive function.

2. User experience of tutors who deliver formative feedback through the use of an end-of-session report

In addressing the research question “What is the user experience of tutors who deliver feedback through an online end-of-session report?”, two main themes emerged from the analysis that explained how participants provided formative feedback for learners. These categories were: (1) *Perceived Purpose* of Feedback Forms, which captured participants’ interpretations of why such forms were used within their institutions, and (2) *Perceived Usability*, which captured whether participants found the forms to be usable, satisfactory, and efficient. The following sections present each theme in detail, supported by participants’ narratives.

Perceived Purpose

The findings revealed three distinct purposes of feedback forms according to tutors: 1) fulfilling administrative requirements, 2) supporting tutors’ work-related functions, and 3) fostering learner development. The results showed that tutors predominantly perceived feedback forms as fulfilling administrative requirements (cited by 8 participants). They are then seen as supporting tutors themselves in order to aid them in recalling the previous session or in order to hand off a learner to another tutor (cited by 6 participants). Finally, only 4 participants mentioned that the purpose of the feedback form was to foster learner development. Upon further

discussion, 10 participants spontaneously mentioned that they do not believe that learners look at the feedback forms.

Table 4 below shows the Coding scheme for the perceived purpose of feedback forms showing sub-themes and representative participant quotes.

Table 4. Perceived Purpose of Feedback Forms Sub-Themes and Representative Participant Quotes

Theme	Sub-Themes	Example
Perceived Purpose	Administrative requirements	“I understand the need for the admin aspect like tracking the work, etc. But it gets repetitive.” (P05)
	Tutor-related functions	“if I for example take two weeks off, someone else will replace me, so they need to read them and to be sure that they understand the flaws, the qualities, the things we have to work on.” (P14)
	Learner development	“It’s more like a reminder for the learner. They can consult the report and say, “Oh yeah, we did that,” or “I remember that mistake.” (P04)

Perceived Usability

Table 5 below shows the Coding scheme for the perceived usability of feedback forms showing sub-themes, representative participant quotes, codes, and number of participants.

Table 5. Perceived Usability of Feedback Forms Sub-Themes, Example Participant Quotes, Codes, and Number of Participants

Theme	Sub-Themes	Participant Example Quotes	Codes	N
Perceived Usability	Effectiveness	“It feels like I have to give a number, and especially in languages—it’s hard to quantify a 60-minute session. How do you evaluate vocabulary? Is it variety?” (P04)	Ratings	8

Table 5. Continued

Theme	Sub-Themes	Participant Example Quotes	Codes	N
Perceived Usability	Satisfaction	“So actually, if you fill more than two lines here, you have a problem because the student can't read and you can't read it either. So, you can't go past those two lines.” (P11)	Ratings	4
			Short Fields	4
			Redundant Fields	2
	Efficiency	Participants self reported on the time it takes them on average to complete a form for every learner.	Very Short (2 minutes)	3
			Short (5 minutes)	7
			Medium (Between 6 and 10 minutes)	3
			Long (More than 10 minutes and up-to 15)	2

A. Effectiveness

Participants mentioned a lack of effectiveness and accuracy when it came to rating their learners on a scale of 1 to 5 in their linguistic performance (Cited by 8 participants). “How do I define performance in conjugation out of 5? I don’t like this method.” (P02). Difficulty was cited by 2 participants in relation to this step in the feedback form “How do you evaluate vocabulary? Is it variety? Correct usage? Syntax? Does it mean correct application or variety? It’s difficult.”

(P04), “it is a little strange when you're talking of a student syntax to rate him by the day.... I find it difficult.” (P11). Other participants described the process as “not very precise” (P01), “a little strange” (P11), and “a bit confusing” (P12).

Participant 13 said “I don't know what it means.” when referring to the numbers and added “sometimes when someone is struggling with something, I don't know if this actually helps.” Another participant cited the contradictory nature of the form in having an open field to write feedback related to the linguistic element then rate someone on that same element. “Sometimes the learner can use good vocabulary but still make mistakes, you know?” (P14). Another participant suggests that the ratings related to vocabulary could be standardized by identifying if the learner “used idioms” and “enough vocabulary” (P15). Participants reported that using a 1-to-5 scale to assess linguistic performance was unclear, inconsistent, and hard to apply, raising concerns about its accuracy and usefulness.

A. Satisfaction

When it came to expressing dissatisfaction, some participants mentioned that they were frustrated with the ratings but they did not express concerns about accuracy, rather, they said that they did not prefer the ratings (Cited by 3 participants). Participants also expressed dissatisfaction with the limited length of the field in which they could type open-ended feedback (cited by 4 participants). In addition to ratings and short fields, participants expressed dissatisfaction regarding the redundancy of certain elements on the feedback form (cited by 2 participants).

When expressing dissatisfaction related to ratings, participants used verbatim like, “Yeah, I don't prefer it, especially for motivation.” (P14), “I don't like to rate people.” (P07), “I would not rate them because why should we evaluate everything formally?” (P11). These types of

responses, while related to the ratings, did not fall under the usability attribute “effectiveness” because participants were sharing their level of comfort as opposed to their thoughts related to the accuracy of ratings.

Regarding participant dissatisfaction related to the short text field available to write open-ended feedback, participants said, (1) “If I put a lot of information in the vocabulary field, there’s a tiny arrow on the side to scroll up and down, but it’s tedious. It’s not immediately visible or presentable.” (P04), (2) “Sometimes we write a lot and it’s hard to format.” (P09), (3) “if you feel more than two lines here, you have a problem because the student can't read and you can't read it either.” (P11), (4) “Sometimes I have so much to say because simply we worked on so many things at the same time during one session that it's not enough. And just having basically one box for me is not enough to have a good reflection on the good and the bad.” (P13).

In addition to ratings and short fields, participants expressed dissatisfaction regarding the redundancy of certain elements on the feedback form (cited by 3 participants). A participant mentioned, “I find it much easier and more efficient to fill out the section “Topic” and mention (in the same field) that we integrated present perfect into the following role play. It’s sometimes a bit tedious to split it up into these sections (linguistic elements)” (P04). Another participant said, “I don't really teach one without the other. I'm not only coming to the lesson teaching pronunciation, for me all of them are always being taught” (P08). Overall, participants mainly expressed discomfort with giving ratings, frustration with the very short open-ended text field and with redundant sections in the feedback form. Their comments focused on personal preference and usability issues, such as limited space to write and repetitive categories, rather than concerns about rating accuracy.

B. Efficiency

Participants were asked how much time it usually took them to complete the ESR and their responses were based on recall and estimation. Three participants completed the form in 2 minutes, seven participants completed it in 5 minutes, three participants completed it between 6 and 10 minutes, and two participants completed it in more than 10 minutes.

3. Tutor perceptions, experiences, concerns, and perceived benefits of AI

A. AI Usage

Participants were asked whether they use AI for tutoring or not and the results show that 6 participants used AI for tutoring while 8 participants did not use AI for tutoring (some of those participants use AI for different situations not related to the scope of this study). Participants who used AI for their tutoring sessions mentioned that it helped them come up with topic ideas, warm-up ideas, sentences using a specific tense, and polish their tone in emails.

B. AI Concerns

When participants were asked about their thoughts regarding AI, spontaneous responses emerged regarding their concerns. Six themes were developed from the content of the transcripts. Table 6 below summarizes the themes, definitions of themes, representative participant quotes, and the number (n) of participants who have mentioned the related concern.

Table 6. Themes, Definitions, and Examples Quotes, # of Participants

Themes	Definition	Example quote	n
Output Accuracy and Nuance	Output accuracy refers to the closeness of a measured value to the true value, while nuance involves subtle distinctions and complexities in data interpretation that affect accuracy and utility (Haegemans et al., 2016).	“But what if a student was sarcastic when communicating ideas?” (P10)	9

Table 6. Continued

Themes	Definition	Example quote	n
Privacy and Security	Privacy is the ability of individuals or groups to control how, when, and to what extent their personally identifiable information is shared, whereas security pertains to the protection of systems and data from threats and unauthorized access (Bishop, 2003).	"I don't think I'd feel right about that. Sometimes they tell me personal things." (P07)	6
Human-Touch	In AI, "human touch" refers to the uniquely human ability for empathy and nuanced understanding that AI cannot fully replicate (Tjernström & Åkerblom-Andersson, 2024)	"I think there's an experience of being with another person that I don't think AI is ever going to recreate" (P08)	6
Panopticon Effect	The Panopticon effect refers to the psychological impact of being observed, leading individuals to regulate their behavior due to the awareness of potential surveillance (Foucault, 1997).	"I know it's not a person, of course, but it's like someone is watching us... it's like a Big Brother kind of thing." (P01)	3
Cognitive Offloading	Cognitive offloading refers to the practice of using external tools or resources to reduce mental effort and enhance cognitive performance, such as writing down information or storing it digitally (Risko & Gilbert, 2016).	"Sometimes you ask students to brainstorm before class, and they come with all these perfect sentences. I'm like, 'OK, you didn't brainstorm. You cheated.'" (P05)	3

C. AI Concerns and AI Usage

AI Concerns were analyzed in relation to a participant's AI usage or lack thereof to identify potential correlations. The sub-theme with the highest mentions was Output Accuracy and Nuance (9 participants). The results showed that participants who used AI for tutoring were more concerned with the correctness and sophistication of the outputs (6 participants) than participants who did not.

The second most coded themes were "privacy and security" and "human touch". Both were mentioned by an equal number of participants (6 participants) and both themes displayed a

similar pattern whereby non-users of AI cited these 2 concerns more frequently than AI users (4 Non-AI Users versus 2 AI Users). The third most coded theme was Unemployment Anxiety, whereby 4 participants mentioned that AI was taking over certain jobs. An equal number of AI users (2) versus non-AI users (2) mentioned this. The fourth most coded themes were “Cognitive Offloading” and “Panopticon Effect” (3 participants each). Although both themes were mentioned by an equal number of participants, they differed in that only non-AI users (3 participants) were concerned with cognitive offloading, and 2 AI users noted the panopticon effect compared to 1 non-AI user. Table 7 below summarizes which AI concerns were prevalent across 2 groups of participants: Participants who used AI for tutoring and those who did not.

Table 7. AI Usage and Themes

AI Usage	Themes					
	Output Accuracy and Nuance	Unemployment Anxiety	Privacy and Security	Human Touch	Cognitive Offloading	Panopticon Effect
Used AI for Tutoring	6	2	2	2	0	2
Didn't Use AI for Tutoring	3	2	4	4	3	1

The values represent the number of participants assigned to each theme.

D. AI Concerns and Technology Knowledge

The means of all technology-related scales (TK, TCK, TPK) were extracted, then quartiles were created using the median-of-halves (Tukey) method in order to find participants that were on the high/medium/low on the quartile related to the technology dimension. The analysis of levels of combined technological dimensions in relation to AI concerns showed that “Output Accuracy and Nuance” was the most mentioned theme regardless of the quartile

participants fell in. Output accuracy and nuance were mentioned by 7 participants in the high to medium quartile but by only 2 participants in the low quartile. “Privacy and Security” was mentioned by 3 participants in the high and medium quartiles and by 3 participants in the low quartile.

“Unemployment Anxiety”, “Human Touch”, and “Panopticon Effect” were concepts that were more familiar to participants with medium to high technology knowledge. “Human Touch” was mentioned by 6 participants in the high and medium quartiles but by none in the low quartile. The “Panopticon Effect” was mentioned by 3 participants in the medium to high quartiles and by none in the low quartile. Low-technology quartile participants were largely absent from references to themes like “Unemployment Anxiety”, “Human Touch”, and the “Panopticon Effect”. Table 8 below summarizes AI concerns and technology knowledge quartiles.

Table 8. AI Concerns and Technology Knowledge

Technology Quartile	Output Accuracy and Nuance	Unemployment Anxiety	Cognitive Offloading	Human Touch	Panopticon Effect	Privacy and Security
High	3	2	1	3	1	1
Medium	4	2	1	3	2	2
Low	2	0	1	0	0	3

Table 9 below shows the intersection of technology knowledge quartiles and AI Usage.

Table 9. AI Concerns related to AI Usage and Technology Knowledge

TK Quartile Themes	High		Medium		Low	
	Uses AI	Doesn't Use AI	Uses AI	Doesn't Use AI	Uses AI	Doesn't Use AI
Output Accuracy and Nuance	1	2	3	1	2	0
Unemployment Anxiety	1	1	1	1	0	0

Table 9. Continued

TK Quartile Themes	High		Medium		Low	
	Uses AI	Doesn't Use AI	Uses AI	Doesn't Use AI	Uses AI	Doesn't Use AI
Cognitive Offloading	0	1	0	1	0	1
Human Touch	1	2	1	2	0	0
Privacy and Security	1	0	0	2	1	2
Total	2	2	3	2	2	2

E. Technology Knowledge and Age

When comparing the results of the technology domain (in the TPACK questionnaire) to age groups, the results showed that 0 participants in the 20s-30s group rated themselves high on the technological domain in comparison to the other age groups. Two of the 4 participants from the mid-life (50s) age group rated themselves as high on the technology domain, and almost half (2 of 5) from the older adults (60+) age group also rated themselves as high on the technological domain. 4 participants (evenly split between young adults and older adults) fell on the low technology domain scale while none of the mid-life adults (50s) fell in this section. Table 10 below summarizes participant technological domain quartiles across age groups.

Table 10. Technology Knowledge Quartiles and Age Groups

Technology Knowledge Quartiles	20s-30s	50s	60-70s	N by Quartile
High	0	2	2	4
Medium	4	2	1	7
Low	2	0	2	4
N by Age	6	4	5	

F. AI Usage and Technology Knowledge

AI usage was compared to the three age groups in this study. Four out of 6 young adults (20s-30s) used AI for tutoring, zero out of 4 midlife adults (50s) in this sample did not use AI for tutoring, and 3 out of 5 older adults (60s-70s) used AI for tutoring. Table 11 below summarizes the age groups, participants in those age groups who used AI for tutoring, and participants who did not.

Table 11. AI Usage in relation to Age

Age Groups	Used AI for Tutoring	Didn't use AI for Tutoring
Young adults (20s–30s)	4	2
Midlife (50s)	0	4
Older adults (60s–70s)	3	2

Four out of 6 participants in the young adults age group used AI for tutoring. For mid-life adults, all 4 did not use AI for tutoring. There was a clear indication that participants in the mid-life age group did not prefer to use AI for tutoring. In the older adults age group, almost half (3 of 5 participants) preferred to use AI for tutoring. Adoption was high among younger and older participants and low among mid-life adults.

V. Discussion

Summary of Results

The findings suggest that tutors predominantly aligned their definition of feedback with cognitive and motivational function. In contrast, feedback aimed at fostering learner self-reflection (a key aspect of the metacognitive function) appeared to be less emphasized in the context of second language tutoring. With regards to how tutors perceive the purpose of the feedback form, it was found that tutors largely interpreted them as administrative aids rather than tools designed to support student learning. Many relied on them to document previous sessions or facilitate transitions between tutors, whereas relatively few linked the forms to promoting

learner growth. Collectively, these insights point to a disconnect between how feedback tools are meant to function and how they are actually perceived by tutors. Following up on the concerns specific to the feedback form, it was found that tutors view rating learners' progress on a scale of 1 to 5, noting that it fails to reflect the complexity of individual learning trajectories.

Another finding from the analysis showed that worries about AI, especially its precision and ability to capture nuance, were the most commonly raised concerns. Participants who incorporated AI into their tutoring were more attentive to the accuracy and sophistication of its responses than those who did not, indicating that direct experience with AI may heighten expectations for output quality. A noticeable pattern emerged in relation to age: individuals in the mid-life group were less inclined to rely on AI for tutoring compared with both younger and older participants.

1. Tutors definition of feedback in the context of asynchronous second language tutoring

Contribution #1: Definition of Feedback

To explore the functions of feedback in asynchronous language tutoring, tutors were asked to describe what feedback meant to them. The findings indicate that cognitive and motivational functions were the most prominent aspects of feedback. This extends Narciss' (2008) feedback framework by showing that, in practice, tutors prioritize supporting learners' immediate understanding and motivation, while metacognitive feedback, which helps learners reflect on how they think and learn, receives comparatively little emphasis. The apparent lower importance of metacognitive feedback in second language tutoring is a noteworthy finding, suggesting that feedback practices are shaped not only by theory but also by the practical demands of tutoring.

2. User experience of tutors who deliver formative feedback through the use of an end-of-session report

Contribution #2: Perceived Purpose of ESR (Feedback Forms)

The findings of this study highlight a tension between the two main purposes of feedback identified by Sadler (2010): promoting student learning and fulfilling administrative requirements. Many tutors described feedback forms as primarily fulfilling administrative purposes such as tracking lesson progress and providing evidence of tutoring, reinforcing Sadler's (2010) observation that feedback systems can shift from formative to administrative purposes. The study also uncovers a third, tutor-focused role for feedback, demonstrating that tutors used forms to support their own teaching processes, such as recalling past sessions and planning future instruction. Overall, these findings broaden the understanding of feedback, showing that online feedback systems can simultaneously fulfill learner growth, administrative functions, and tutor-supportive functions.

Contribution #3: Reliability and Challenges of Numeric Rating Scales

Tutors reported that rating learners on their performance from a scale of 1 to 5 was not effective. The results suggest that numeric rating scales in language assessment can limit meaningful feedback, contributing to literature by Lumley (2002), Sadler (2010), and Hyland (2000) by showing that rating language skills using numbers oversimplifies learners' performance and reduces their ability to act on feedback. While numeric ratings may make assessments appear standardized, tutors reported difficulty in selecting appropriate numbers. This indicates that when feedback is treated as a one-way transmission of information, it can fail to guide learners in interpreting and regulating their own learning. Furthermore, numeric scales also impacted the practical use of feedback forms, as tutors were often unable to complete them

within the expected two-minute timeframe, suggesting that the design of numeric ratings can interfere with both clarity and efficiency in feedback practice.

3. Tutor perceptions, experiences, concerns, and perceived benefits of AI

Contribution #4: AI Usage and Age

Unexpected patterns in the relationship between age and AI adoption among the 15 tutors in this study were found. The first pattern emerged in AI tutoring usage, showing that there was high adoption in older and younger tutors but not so much in middle-aged tutors. The high adoption among older adult tutors is particularly striking given the cautious stance toward AI often attributed to this demographic. This pattern contradicts the linear relationship between age and technology adoption often reported in the literature, where younger users are expected to demonstrate consistently higher adoption rates such as Lāma and Lastovska's (2025) findings which show that participants under age 49 demonstrated higher self-rated AI competence. The tutors in this study may represent a self-selected group with higher pedagogical innovation orientation, or they may perceive AI tutoring tools as practical solutions to specific organizational challenges. These results suggest that age-related differences in AI adoption may vary across contexts, offering a nuanced perspective to ongoing debates about AI readiness in education.

Contribution #5: AI Concerns and AI Usage

The second pattern emerged in tutors' AI usage, showing that a person's engagement with AI shapes both the type and intensity of their concerns. Although fears of job displacement appeared across users and non-users, non-users focused more on risks related to cognitive offloading, whereas active users were more attentive to surveillance and data-use issues. This contributes to the literature (e.g., Bewersdorff et al., 2025) by showing that hands-on AI

experience shifts users' attention toward output accuracy and transparency, while limited experience amplifies concerns around privacy, security, and the loss of human interaction.

Contribution #6: AI Concerns and Technology Knowledge

Results indicate that self-rated technological knowledge affects AI-related concerns in more nuanced ways than anticipated, contributing to the literature (Pan et al., 2025; Bewersdorff et al., 2025) by showing that greater self-assessed expertise broadens the variety, depth, and focus of concerns expressed. Participants with higher technology knowledge scores highlighted systemic and social issues, including AI and its potential impact on the job market, maintaining a human touch, and surveillance or data-use challenges, reflecting a wider awareness of AI's societal effects. In contrast, participants with lower technology knowledge tended to concentrate on concerns such as privacy, security, and accuracy of AI outputs. Worries about AI accuracy were common across all levels, and mentions of cognitive offloading were rare and evenly spread, suggesting that some concerns are widely shared regardless of self-perceived knowledge. Overall, the findings suggest that increased technological knowledge does not simply lessen apprehension; instead, it shapes the nature and sophistication of AI concerns.

Contribution #7: Age and Technology Knowledge

Analysis of TPACK technology domain scores revealed distinct age-related patterns, specifically, no young adult participants rated themselves high in technological competence, while 50% of midlife adults and 40% of older adults provided high self-ratings. Several explanations may account for this unexpected result. Young adult tutors, having grown up as digital natives, may possess more sophisticated understanding of technological complexity and therefore recognize gaps in their own knowledge, a pattern consistent with the Dunning-Kruger effect (1999) where greater actual competence leads to more modest self-assessments.

Conversely, older and midlife adults may evaluate themselves relative to age-matched peers rather than younger cohorts, rating themselves highly based on personal growth trajectories and comparison to their own starting points.

Contribution #8: AI Usage and Technology Knowledge

Results suggest that midlife adults' high self-rated technology competence does not predict their AI usage for tutoring, which contributes to the literature (Lāma & Lastovska, 2025) by showing that factors such as curiosity, openness to experimentation, and pedagogical needs may drive AI use independently of perceived technology knowledge or skill. In this study, young adults with lower self-rated competence demonstrated the highest AI usage rate for tutoring (67%), while midlife adults, despite high self-ratings, did not use AI for tutoring. This finding contributes to the literature by highlighting that self-rated technology competence is influenced by relative self-assessment, not just actual ability, which has implications for research on technology adoption, training, and AI readiness.

Design and Practical Implications

This section explains the main design lessons from the findings and how they can help shape more useful and user-friendly feedback tools for tutors. Based on our findings, it is recommended to remove numeric rating scales from ESRs or feedback forms in a language learning context, as they may fail to provide meaningful support for learner development. It is also recommended to replace the linguistic elements section and replace it with a large text field. It is also recommended to utilize AI to summarize sessions to fulfill the administrative, tutor-related, and learner support functions. Taken together, these implications point to several ways feedback tools can be reshaped to reflect how tutors actually work, address the challenges they face, and enhance both ease of use and educational value to learners.

Limitations and Research Avenues

This study was limited to participants from a single company. As a result, the findings may not have fully represented the experiences of tutors in other contexts. The sample size only shows trends which are typical of quantitative analyses and the patterns found in this study may reflect unique characteristics of this particular group of tutors rather than the trends of the broader population . Future studies could replicate the study with larger sample sizes to increase the robustness and transferability of the results.

Another key limitation of the study was related to the usage of the TPACK questionnaire to assess participants' technological, pedagogical, and content knowledge. The questionnaire relies on participants' self-assessment of these dimensions. Self-reported measures could be influenced by factors such as overconfidence, underestimation, or differing interpretations of the questionnaire items, which may have affected the accuracy and reliability of the data. In future studies, other assessments to determine TPACK could be utilized such as performance based TPACK assessments (Rosenberg et al., 2015).

Regarding the data analysis method in this study, which involved coding respondents' verbatim statements according to the three usability attributes defined by ISO 9241-11 (2018) which are effectiveness, satisfaction, and efficiency, Doi (2021) noted that some participants may struggle to verbalize their thoughts, and that the complexity and duration of the task can influence both the content and quantity of their verbal responses.

On another note, adopting Narciss's (2008) framework directed the analysis toward the cognitive, motivational, and metacognitive functions of feedback. Applying a different conceptual model (e.g., Hattie & Timperley, 2007; Shute, 2008) could have shifted the focus

toward other dimensions, such as the level, format, or timing of feedback, which may have resulted in alternative interpretations of feedback quality and effectiveness.

Another limitation of this study is the uneven age distribution within the sample. Because there were no participants aged 20 to 28 and 38 to 50, the dataset has sizable gaps that limit its ability to reflect the broader adult population. The resulting age categories: 29 to 37, 51 to 58, and 61 to 77 also cover different ranges, with the oldest group spanning a particularly wide set of life stages. These missing segments create large jumps between groups rather than a smooth progression across ages, making it difficult to capture gradual changes or trends. Consequently, the age-related insights drawn from this sample have limited generalizability and may not apply to younger adults, middle-aged individuals, or other unrepresented groups.

Concluding Remarks

This study shows that tutors' feedback practices in asynchronous online environments are shaped by pedagogy, practical constraints, institutional priorities, and personal experience with technology. Furthermore, feedback acts as a multifaceted process serving administrative functions, learner-developmental, and tutor supportive roles. The results of this study also show that AI adoption relies more on curiosity and pedagogical relevance than age or technological skill. This insight highlights that effective feedback, AI-assisted or not, relies on human connection, reflection, and adaptability to foster successful learning experiences.

References

- Alnemrat A, Aldamen H, Almashour M, Al-Deaibes M and AlSharefeen R (2025) AI vs. teacher feedback on EFL argumentative writing: a quantitative study. *Front. Educ.*
<https://doi.org/10.3389/feduc.2025.1614673>
- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55(5), 469–480.
<https://doi.org/10.1037/0003-066X.55.5.469>
- Alshahrani, A., & Mostafa, A. M. (2025). *Enhancing the use of artificial intelligence in architectural education – case study Saudi Arabia*. *Frontiers in Built Environment*, 11, Article 1610709. <https://doi.org/10.3389/fbuil.2025.1610709>
- Bell, L. (2019). Examining tutoring strategies in asynchronous screencast tutorials. *Global Society for Online Literacy Educators*. <https://gsole.org/olor/role/vol2.iss2.f>
- Bewersdorff, A., Hornberger, M., Nerdel, C., & Schiff, D. S. (2025). AI advocates and cautious critics: How AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy. *Computers and Education: Artificial Intelligence*, 8, 100340.
<https://doi.org/10.1016/j.caeai.2024.100340>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Bressler, M., & Von Bergen, C. (2014). The sandwich feedback method: Not very tasty. *Journal of Behavioral Studies in Business*, 7. Retrieved from
<https://www.aabri.com/manuscripts/141831.pdf>

- Chai, C. S., Koh, J. H. L., & Tsai, C.-C. (2013). A Review of Technological Pedagogical Content Knowledge. *Journal of Educational Technology & Society*, 16(2), 31-51.
<http://www.jstor.org/stable/jeductechsoci.16.2.31>
- Chan, C.K.Y., Hu, W. Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *Int J Educ Technol High Educ* 20, 43 (2023).
<https://doi.org/10.1186/s41239-023-00411-8>
- Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19(2), 189–211. <https://doi.org/10.2307/249688>
- Dohrenwend, A. (2002). Serving up the feedback sandwich. *Family Practice Management*, 9(10), 43–46. <https://pubmed.ncbi.nlm.nih.gov/12469676/>
- Doi, T. (2021). Usability Textual Data Analysis: A Formulaic Coding Think-Aloud Protocol Method for Usability Evaluation. *Applied Sciences*, 11(15), 7047.
<https://doi.org/10.3390/app11157047>
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. MIT Press.
<https://doi.org/10.7551/mitpress/5657.001.0001>
- Foucault, M. (1977). *Discipline and punish: The birth of the prison*. Pantheon Books.
- Forster, M., Weiser, C., & Maur, A. (2018). How feedback provided by voluntary electronic quizzes affects learning outcomes of university students in large classes. *Computers & Education*, 121, 100–114. <https://doi.org/10.1016/j.compedu.2018.02.012>
- Guo, K & Pan, M & Li, Y & Lai, C. (2024). Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The Internet and Higher Education*. 63. <https://doi.org/10.1016/j.iheduc.2024.100962>.

- van den Haak, M., De Jong, M., & Jan Schellens, P. (2003). Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue. *Behaviour & Information Technology*, 22(5), 339–351. <https://doi.org/10.1080/0044929031000>
- Hyland, F. (2000). ESL writers and feedback: giving more autonomy to students. *Language Teaching Research*, 4(1), 33-54. <https://doi.org/10.1177/136216880000400103>
- Haegemans, T., Snoeck, M., & Lemahieu, W. (2016). Towards a precise definition of data accuracy and a justification for its measure. In *Proceedings of the International Conference on Information Quality*. Retrieved from <https://www.semanticscholar.org/paper/Towards-a-Precise-Definition-of-Data-Accuracy-and-a-Haegemans-Snoeck/0cae60d5683fc3b4ba63b82a5847e6f047511156>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., et al. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32(3), 504–526. <https://doi.org/10.1007/s40593-021-00239-1>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- International Organization for Standardization. (2018). ISO 9241-11:2018 Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. ISO. <https://www.iso.org/standard/63500.html>
- Jensen, L., Bearman, M., & Boud, D. (2021). Understanding feedback in online learning – A critical review and metaphor analysis. *Computers & Education*, 173, 104271. <https://doi.org/10.1016/j.compedu.2021.104271>

- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Lāma, G., & Lastovska, A. (2025). AI competence and sentiment: A mixed-methods study of attitudes and open-ended reflections. *Frontiers in Artificial Intelligence*, 8. <https://doi.org/10.3389/frai.2025.1658791>
- Lumley, T. (2002). Assessment criteria in a large-scale writing test: what do they really mean to the raters? *Language Testing*, 19(3), 246-276. <https://doi.org/10.1191/0265532202lt230oa>
- Lilly, J., Richter, U., & Rivera-Macias, B. (2010). Using feedback to promote learning: Student and tutor perspectives. *Practitioner Research in Higher Education*, 4. <https://eric.ed.gov/?id=EJ1130656>
- Li, J., & Huang, J.-S. (2020). *Dimensions of artificial intelligence anxiety based on the integrated fear acquisition theory*. *Technology in Society*, 63, 101410. <https://doi.org/10.1016/j.techsoc.2020.101410>
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, 19(1), 37–66. <https://doi.org/10.1017/S0272263197001034>
- Mandouit, L., & Hattie, J. (2023). Revisiting “The power of feedback” from the perspective of the learner. *Learning and Instruction*, 84, 101718. <https://doi.org/10.1016/j.learninstruc.2022.101718>
- Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>

- Mishra, P., Warr, M., & Islam, R. (2023). TPACK in the age of ChatGPT and Generative AI. *Journal of Digital Learning in Teacher Education*, 39(4), 235–251.
<https://doi.org/10.1080/21532974.2023.2247480>
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(1), 9.
<https://doi.org/10.1186/s40561-024-00295-9>
- Maier, U., Wolf, N., & Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education*, 95, 85–98. <https://doi.org/10.1016/j.compedu.2015.12.002>
- Metcalf, J., Kornell, N., & Finn, B. (2009). Delayed versus immediate feedback in children's and adults' vocabulary learning. *Memory & Cognition*, 37(8), 1077–1087.
<https://doi.org/10.3758/MC.37.8.1077>
- Moss, C. M., & Brookhart, S. M. (2019). Advancing formative assessment in every classroom: A guide for instructional leaders (2nd ed.). ASCD.
<https://www.ascd.org/books/advancing-formative-assessment-in-every-classroom-a-guide-for-instructional-leaders-2nd-edition?variant=120005>
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 125–144). Lawrence Erlbaum Associates/Taylor & Francis.
https://www.researchgate.net/publication/233833269_Feedback_Strategies_for_Interactive_Learning_Tasks

- Pajares, M. F. (1992). Teachers' Beliefs and Educational Research: Cleaning Up a Messy Construct. *Review of Educational Research*, 62(3), 307-332.
<https://doi.org/10.3102/00346543062003307>
- Pan, Z., Moore, O. A., Papadimitriou, A., & Zhu, J. (2025). AI literacy and trust: A multi-method study of human–GAI team collaboration. *Computers in Human Behavior: Artificial Humans*, 4, 100162. <https://doi.org/10.1016/j.chbah.2025.100162>
- Prochazka, J., Ovcari, M., & Durinik, M. (2020). Sandwich feedback: The empirical evidence of its effectiveness. *Learning and Motivation*, 71, 101649.
<https://doi.org/10.1016/j.lmot.2020.101649>
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>
- Rosenberg, J. M., Greenhalgh, S. P., & Koehler, M. J. (2015). *Performance assessment of teachers' TPACK using artifacts from digital portfolios*. In L. Liu & D. C. Gibson (Eds.), *Research highlights in technology and teacher education 2015* (pp. 189–196). Society for Information Technology & Teacher Education (SITE). Retrieved from
https://www.matt-koehler.com/publications/rosenberg_et_al_2015.pdf
- Said, N., Potinteu, A. E., Brich, I., Buder, J., Schumm, H., & Huff, M. (2023). An artificial intelligence perspective: How knowledge and confidence shape risk and benefit perception. *Computers in Human Behavior*, 149, 107855.
<https://doi.org/10.1016/j.chb.2023.107855>
- Schmid, M., Brianza, E., & Petko, D. (2020). Developing a short assessment instrument for Technological Pedagogical Content Knowledge (TPACK.xs) and comparing the factor

- structure of an integrative and a transformative model. *Computers & Education*, 157, 103967. <https://doi.org/10.1016/j.compedu.2020.103967>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144. <https://doi.org/10.1007/BF00117714>
- Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5), 535–550. <https://doi.org/10.1080/02602930903541015>
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*. 78. 153-189. <https://doi.org/10.3102/0034654307313795>
- Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29(7), 4-14. <https://doi.org/10.3102/0013189X029007004>
- Tjernström, L., & Åkerblom-Andersson, C. (2024). *It's smart, but you know, it lacks that human touch! Exploring and designing for dynamic user control in AI-driven automated systems* (Master's thesis). Umeå University. <https://umu.diva-portal.org/smash/get/diva2:1876242/FULLTEXT01.pdf>
- Trumbull, E., & Lash, A. (2013). *Understanding Formative Assessment: Insights from Learning Theory and Measurement Theory*. WestEd. <https://www.wested.org/resource/understanding-formative-assessment-insights-from-learning-theory-and-measurement-theory>
- Tortella, F., Palese, A., Turolla, A., Castellini, G., Pillastrini, P., Landuzzi, M. G., & Rossettini, G. (2025). Knowledge and use, perceptions of benefits and limitations of artificial intelligence chatbots among Italian physiotherapy students: A cross-sectional national study. *BMC Medical Education*, 25(1), 572. <https://doi.org/10.1186/s12909-025-07176-w>

- Wang, N., & Li, M. (2024). Teachers' perceptions of the risks and benefits of AI in higher education: A case study of ERNIE Bot. *Innovations in Education and Teaching International*, 1–13. <https://doi.org/10.1080/14703297.2024.2432429>
- Walker, B. (2018). A Defining Moment for Personal Tutoring: Reflections on Personal Tutor Definitions and their Implications Ben W. Walker. <https://e-space.mmu.ac.uk/626048/>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage. <https://uk.sagepub.com/en-gb/eur/case-study-research-and-applications/book250150>
- Yoon, S. Y., Miszoglad, E., & Pierce, L. R. (2023). Evaluation of ChatGPT feedback on ELL writers' coherence and cohesion. *arXiv*. <https://arxiv.org/abs/2310.06505>
- Zhai, X., & Nehm, R. H. (2023). AI and formative assessment: The train has left the station. *Journal of Research in Science Teaching*, 60(6), 1390–1398. <https://doi.org/10.1002/tea.21885>
- Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102. <https://doi.org/10.1016/j.asw.2018.02.004>

Appendices

Appendix A: ESR

1. Learner Information / Information de l'étudiant	
Student / Apprenant :	esi_feb28 test
Tutor / Tuteur :	Professiorri
Lesson Date :	05/29/2025 08:30
Completed Date :	05/29/2025 14:49
Module :	1
Total support per Module :	21 hrs
Session Duration / Durée Session :	60 mins
Total Sessions :	21
Current Session # :	1
Remaining Sessions :	20
Additional non-module based homework / Temps d'autoapprentissage hors du module :	
Time spent in self-study since last session / Temps auto-apprentissage depuis session précédente (hrs) :	N/A
Online module-based homework / Temps d'autoapprentissage dans le module :	
Time spent in self-study since last session / Temps auto-apprentissage depuis session précédente :	1.41 hrs
Module Progress / Progres dans le module :	0

2. Pedagogical content of this lesson / Résumé pédagogique de cette leçon	
A. Linguistic Elements / Éléments linguistiques : (Must complete at least one)	
Conjugation / Conjugaison :	
Syntax / Syntaxe :	
Vocabulary / Vocabulaire :	
Spelling / Orthographe :	
Pronunciation / Prononciation :	
Comprehension / Compréhension :	
B. Dialogue / Role play / Jeu de rôles :	
Topic / Thème :	

3. Learner Performance / Performance de l'apprenant (*This section is required prior to saving)	
Conjugation / Conjugaison :	☒ N/A ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
Syntax / Syntaxe :	☒ N/A ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
Vocabulary / Vocabulaire :	☒ N/A ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
Spelling / Orthographe :	☒ N/A ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
Fluency / Débit :	☒ N/A ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
Pronunciation / Prononciation :	☒ N/A ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
Participation / Motivation :	☒ N/A ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+
Comprehension / Compréhension :	☒ N/A ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+

4. Learner Follow-Up Form - Rapport de suivi de l'apprenant	
A. Homework / Specific areas to review / Devoirs / Points particuliers à réviser : (*This section is required prior to saving)	
B. Recommended links / Liens suggérés :	

Appendix B: Interview Guide

Part 1

I would like to inquire about feedback within the scope of your role as a tutor.

1.1 Firstly, how would you define feedback?

1.2 What elements does it consist of?

1.3 What do you think makes feedback useful or meaningful to learners?

1.4 Can you describe your general approach to giving feedback to learners?

1.5 How often do you find that you provide personalized feedback or more structured feedback across all learners?

1.6 How come?

1.7 Is it because you don't have the opportunities/time/don't perceive as useful? Are there learning options where you don't provide this modality.

1.8 How do you think your feedback influences a student's learning and development?

Part 2

Now we will transition from general feedback to a specific form of feedback: the end-of-session report completed after each tutoring session in the program.

2.1 What was your experience like while filling out the end-of-session report to a learner/group?

- When and how often do you fill out the end-of-session report?
- What are the steps required to fill out the report?
- What is the time and effort needed to fill out a report?

2.2 (Show an example of an empty ESR report) Here is an example of an end-of-session report.

Could you walk me through this report and explain the different sections to me? Tell me:

- How would you fill it out?

- Overall, how effective do you find this report to be?
(Probes: Organization, visual presentation, content, interaction)
- Are there aspects of this report that are challenging?
- What changes might make this report more effective for tutors to use?

2.3 What do you consider the most important part of the ESR?

- What do you consider the least important part of the ESR?

2.4 Do you consult previous ESRs when:

- You start tutoring a new learner?
- With an actual student, when crafting new reports?

2.5 Do you reference any other materials, such as your personal notes, when completing the ESR?

- How do your notes and the report complement or compete with each other?

2.6 Do you fill out the ESR differently based on what module the student is at?

2.6 If there was no end-of-session report as you currently know it, how would you approach providing feedback to students?

2.8 What's something you wish you knew before you started as a tutor regarding filling out the end-of-session report in particular?

Part 3

I will now stop sharing my screen and move to the last part of the interview. I would like to discuss your ideas about AI use in the context of your work as a tutor.

3.1 Do you use AI tools (such as ChatGPT)? What other AI tools do you use?

3.2 Have you used it specifically for providing feedback to students, whether it be performance feedback, a summary of the session, or guidance on homework?

- (if yes) How would you describe your experience with AI?
- (if yes) To what extent and in what aspect was it useful?

3.3 Do you think AI can play a role in crafting the end-of-session report for learners?

- (If yes) How would you perceive its role? In what context or situation could you see AI being useful/beneficial?
 - How would you compare the potential usage of AI in those 3 contexts: performance feedback, session's summary and guidance for homework?
- If not, why? Where do you think AI could have limitations?
 - How would you compare the limitations in those 3 contexts: performance feedback, session's summary and guidance for homework?

3.4 How would you feel about an AI tool used to analyze sessions with students to generate end-of-session Reports?

- Is there anything you like or dislike about it?
- Do you foresee any benefits?
- Do you foresee any obstacles?

3.5 Do you think AI can play a role in determining what homework can be provided to a learner?

3.6 Do you think AI can play a role in crafting a summary of the session?

Part 5

5.1 Is there anything else you'd like to share about your experiences or thoughts on feedback, electronic feedback, or AI-generated feedback?

Appendix C: Questionnaire

C1. Demographic Questionnaire

1. Please indicate your age.
2. What sex were you assigned at birth?
 - Male
 - Female
 - Prefer not to say
3. What is your gender?
 - Man
 - Woman
 - Prefer not to answer
4. What is the highest level of education you have completed?
 - Cegep (College)
 - Undergraduate Degree (Bachelor)
 - Undergraduate Certificate
 - Graduate Degree (Master's)
 - Post Graduate Degree (PhD)
5. How long have you been a tutor at the organization?
 - 6 to 12 months
 - 13 to 24 months
 - More than 24 months

C2. TPACK Questionnaire

The following questionnaire was used to assess teachers' Technological, Pedagogical, and Content Knowledge (TPACK). Items were rated on a 7-point Likert scale:

1 = Strongly Disagree

2 = Disagree

3 = Somewhat Disagree

4 = Neutral

5 = Somewhat Agree

6 = Agree

7 = Strongly Agree

PK (Pedagogical Knowledge): Knowledge about teaching methods and practices.

1. Please rate the following statements.

- I can adapt my teaching based upon what students currently understand or do not understand.
- I can adapt my teaching style to different learners.
- I can use a wide range of teaching approaches in a classroom setting.

CK (Content Knowledge): Knowledge about the subject matter.

2. Please rate the following statements based on your experience with providing feedback to learners.

- I have sufficient knowledge about my teaching subject.
- I can use a subject-specific way of thinking in my teaching subject.
- I know the basic theories and concepts of my teaching subject.

TK (Technological Knowledge) Understanding of various technologies.

3. Please rate the following statements regarding digital technologies (computers, tablets, mobile phones, AI, etc.).

- I keep up with important new technologies.
- I frequently play around with technology.
- I know about a lot of different technologies.

PCK (Pedagogical Content Knowledge): How to teach specific content effectively.

4. Please rate the following statements.

- I know how to select effective teaching approaches to guide student thinking and learning in my teaching subject.
- I know how to develop appropriate tasks to promote students' complex thinking of my teaching subject.
- I know how to develop exercises with which students can consolidate their knowledge of my teaching subject.

TPK (Technological Pedagogical Knowledge) How teaching changes when tech is used

5. Please rate the following statements.

- I can choose technologies that enhance the teaching approaches for a lesson.
- I can choose technologies that enhance students' learning for a lesson.
- I can adapt the use of the technologies that I am learning about to different teaching activities.

TCK (Technological Content Knowledge): How technology and content interact.

7. Please rate the following statements based on your experience with writing lesson plans.

- I know how technological developments have changed the field of my subject.
- I can explain which technologies have been used in research in my field.

- I know which new technologies are currently being developed in the field of my subject.

TPCK Integrative knowledge of content, pedagogy, and technology.

8. Please rate the following statements based on your experience with writing lesson plans.

- I can use strategies that combine content, technologies, and teaching approaches that I learned about in my coursework in my classroom.
- I can choose technologies that enhance the content for a lesson.
- I can select technologies to use in my classroom that enhance what I teach, how I teach, and what students learn.

Appendix D: Participant Ages

Age Groups	Age Range	Participant ages	n
Young adults (20s–30s)	29-37	29, 32, 33, 35, 37, 37	6
Midlife (50s)	51-58	51, 54, 55, 58	4
Older adults (60s–70s)	61-77	61, 64, 67, 74, 77	5

Footnotes

This article is in preparation for Language Teaching Research, and it has been formatted accordingly.

Chapter 3

Conclusion

This research examined how tutors understand and deliver feedback in asynchronous online learning environments, as well as how AI influences their teaching experiences and perceptions. By analyzing tutors' views on feedback functions, administrative expectations, and AI integration, the study provides a nuanced understanding of how feedback operates within digital education systems that balance pedagogical goals and institutional demands.

The study found that tutors most often described feedback in terms of its cognitive and motivational purposes. Feedback was viewed primarily as a means to clarify understanding, support learning progress, and encourage continued effort, reflecting the central components identified by Narciss (2008). However, metacognitive feedback, which helps learners evaluate and regulate their own learning, was rarely emphasized. This limited focus suggests that while tutors value feedback that strengthens comprehension and motivation, they may devote less attention to feedback that fosters independent learning and self-reflection. This imbalance highlights an opportunity to explore how metacognitive approaches can be better incorporated into asynchronous tutoring, particularly in language learning contexts where dialogue tends to be delayed or text-based.

A second key finding relates to how tutors perceive the purpose of institutional feedback systems. Although feedback is intended to facilitate reflection and dialogue, many tutors in this study saw it primarily as an administrative requirement, something to record tutoring activity and demonstrate accountability, rather than to improve learning. This observation supports Sadler's (2010) argument that institutional feedback systems often prioritize evaluation over

development. Some tutors questioned whether students even read the feedback provided in these systems, suggesting that such forms may have become procedural rather than pedagogical.

Nonetheless, this study also identified a third function of feedback that extends beyond existing theoretical models: a tutor-centered role. Many tutors reported using feedback forms as practical tools to help them plan lessons, recall learner progress, and support collaboration between colleagues. This finding broadens the understanding of feedback's role in online teaching, suggesting that it supports not only learners and institutions but also tutors themselves. Recognizing this additional function provides an opportunity to design systems that value tutors' reflective and organizational needs alongside student learning outcomes.

The findings also revealed several unexpected patterns related to AI usage, age, and self-rated technological competence. Contrary to common assumptions that younger educators are more likely to use AI, older adult tutors in this sample reported relatively high levels of AI use for tutoring, while midlife tutors, who rated themselves as most competent with technology did not use AI at all. This challenges the assumption of a simple, linear relationship between age and technology adoption often cited in studies such as Lāma and Lastovska (2025). These results suggest that factors like curiosity, problem-solving needs, and willingness to experiment may play a greater role in AI adoption than either age or self-perceived skill level. Moreover, the lower self-assessments of technological ability among younger tutors may indicate a more critical self-awareness, aligning with the Dunning–Kruger effect (1999), where higher competence can correspond with more cautious self-evaluation.

When exploring AI-related attitudes and concerns, distinct differences emerged between users and non-users of AI. Tutors who used AI in their work tended to focus on accuracy, reliability, and nuance of output, reflecting concerns based on direct interaction with the tools. In

contrast, non-users expressed greater concern over privacy, security, and the loss of human connection, likely shaped by broader social or ethical discussions rather than hands-on experience. Interestingly, both groups expressed worries about job security, indicating that fears of AI-driven displacement persist regardless of usage level.

Further analysis revealed that tutors with higher self-rated technological knowledge expressed a broader range of nuanced concerns, including issues related to data surveillance and maintaining the human dimension of teaching. This supports Pan et al.'s (2025) idea of “informed skepticism,” suggesting that technological competence does not necessarily reduce anxiety about AI. Instead, knowledgeable users tend to raise more sophisticated and systemic issues. The results also revealed that privacy concerns remain common across all levels of technological ability, possibly reflecting their prominence in public discourse. Together, these findings illustrate that educators’ concerns about AI are shaped by both personal experience and general awareness, with expertise leading to deeper, not necessarily fewer, reservations.

Implications for Practice and Policy

The study’s findings have several implications for institutions and educators. First, feedback systems in online education should be restructured to serve their formative purpose rather than functioning solely as administrative tools. Systems that allow tutors to customize comments, connect feedback across sessions, and invite learner responses could encourage more meaningful engagement and reflection.

Second, professional development initiatives around AI should recognize that confidence and actual use do not always coincide. Training programs should therefore focus on both the technical and pedagogical aspects of AI use, emphasizing critical evaluation of AI outputs, strategies for maintaining the human element in feedback, and awareness of privacy and ethical

considerations. Open discussion about AI skepticism can also help educators approach new technologies thoughtfully and responsibly rather than fearfully.

Concluding Remarks

In summary, this study enhances understanding of feedback in asynchronous online environments by showing that tutors' practices are influenced not only by pedagogical theory but also by practical constraints, institutional priorities, and personal experiences with technology. It highlights the need to view feedback as a multifaceted process, one that serves administrative, developmental, and tutor-supportive functions.

The research also challenges the assumption that technology adoption follows predictable patterns based on age or perceived competence. Instead, it suggests that openness, curiosity, and pedagogical relevance are more decisive factors in AI integration. By examining how tutors define, deliver, and reflect on feedback, as well as how they navigate the promises and challenges of AI, this study contributes to ongoing discussions about the evolving nature of teaching in digital education. Ultimately, effective feedback, whether supported by AI or not, depends on maintaining the human capacity to connect, reflect, and adapt in ways that promote genuine learning and understanding.

References

- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Lāma, G., & Lastovska, A. (2025). AI competence and sentiment: A mixed-methods study of attitudes and open-ended reflections. *Frontiers in Artificial Intelligence*, 8. <https://doi.org/10.3389/frai.2025.1658791>
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 125–144). Lawrence Erlbaum Associates/Taylor & Francis. https://www.researchgate.net/publication/233833269_Feedback_Strategies_for_Interactive_Learning_Tasks
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 125–144). Lawrence Erlbaum Associates/Taylor & Francis. https://www.researchgate.net/publication/233833269_Feedback_Strategies_for_Interactive_Learning_Tasks
- Pan, Z., Moore, O. A., Papadimitriou, A., and Zhu, J. (2025). AI literacy and trust: a multi-method study of human-GAI team collaboration. *Comput. Hum. Behav.* 4:100162. <https://doi.org/10.1016/j.chbah.2025.100162>

Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal.

Assessment & Evaluation in Higher Education, 35(5), 535–550.

<https://doi.org/10.1080/02602930903541015>

Declaration of AI Use

The authors declare that Gen AI was used in the creation of this manuscript. The authors verify and take full responsibility for the use of generative AI solely for language refinement, clarity improvement, and reference formatting; all the information, analysis, interpretations, and conclusions were developed by the authors.