



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS: ESTUDOS LINGUÍSTICOS E
LITERÁRIOS

Vítor Pluceno Behnck

**An exploratory study of Brazilian undergraduate students' engagement with
ChatGPT's automated writing evaluation**

Florianópolis,
2025

Vítor Pluceno Behnck

**An exploratory study of Brazilian undergraduate students' engagement with
ChatGPT's automated writing evaluation**

Dissertação submetida ao Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Mestre em Linguística Aplicada.

Orientadora: Profa. Donesca Cristina Puntel Xhafaj,
Dra.

Florianópolis,
2025

Pluceno Behnck, Vitor

An exploratory study of Brazilian undergraduate students' engagement with ChatGPT's automated writing evaluation / Vitor Pluceno Behnck ; orientadora, Donesca Cristina Puntel Khafaj, 2025.
117 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente, Florianópolis, 2025.

Inclui referências.

1. Inglês. 2. ChatGPT. 3. Aquisição de Segunda Língua. 4. Letramento em IA. 5. Linguística Aplicada. I. Khafaj, Donesca Cristina Puntel . II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente. III. Título.

Vítor Pluceno Behnck

**An exploratory study of Brazilian undergraduate students' engagement with
ChatGPT's automated writing evaluation**

O presente trabalho em nível de Mestrado foi avaliado e aprovado, em 27 de agosto de 2025, pela banca examinadora composta pelos seguintes membros:

Profa. Dra. Kyria Finardi
Universidade Federal do Espírito Santo

Prof. Dr. Leonardo da Silva
Universidade Federal de Santa Catarina

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestre em Linguística Aplicada.



Documento assinado digitalmente
Alinne Balduino Pires Fernandes
Data: 02/10/2025 09:40:56-0300
CPF: ***.588.801-**
Verifique as assinaturas em <https://v.ufsc.br>

Coordenação do Programa de Pós-Graduação



Documento assinado digitalmente
Donesca Cristina Puntel Xhafaj
Data: 01/10/2025 10:30:11-0300
CPF: ***.838.100-**
Verifique as assinaturas em <https://v.ufsc.br>

Profa. Donesca Cristina Puntel Xhafaj, Dra.
Orientadora

Florianópolis, 2025.

ACKNOWLEDGEMENTS

The years during which this thesis was written (early 2023 to mid-2025) were deeply transformative in my life. During this period, I transitioned from being a young researcher to an adult and technical writer while completing my Master's thesis. This work is also, to a certain degree, biographical: it depicts my experience as a researcher on a challenging, complex, and intricate topic, and the strategies I employed to decipher and interpret it.

This work is not about machines. Instead, it is about what people do with them. Therefore, I would like to thank all the people who have supported me throughout this period: my family, friends, professors, and colleagues. A special acknowledgement to Professor Donesca, who believed in me and my ideas when I did not. This thesis is only possible because all of you were there. I am grateful to CAPES for the financial support and to UFSC for providing me with my education during my undergraduate and master's programs.

“You're human, too. You understand love. With that in mind, surely you understand how even the most useless tiny people on this planet are more real than any probabilistic word predictor, and how you must be too. Human obsolescence is not here and can never be.” (Wang, 2023).

ABSTRACT

This exploratory study investigates how Brazilian undergraduate students interact with and incorporate ChatGPT's feedback during an English as a Second Language (L2) writing assignment, as well as how their perceptions of the tool evolve before and after the feedback task. Guided by two research questions, the study explores (1) how students integrate ChatGPT's feedback into their L2 writing and (2) how their perceptions of ChatGPT as an Automated Writing Evaluation (AWE) tool change after engaging with the system. Adopting a cross-sectional, quasi-experimental, posttest-only design, the study involved seven undergraduate students ($N = 7$) from various programs at the Universidade Federal de Santa Catarina (UFSC), all with intermediate English proficiency (B1-B2 CEFR level). Data collection included an initial in-class handwritten essay, a pre-task questionnaire, a one and a half hour training session on ChatGPT (covering functionalities, prompt-writing strategies, and ethical considerations), a revision task using ChatGPT's feedback, and a post-task questionnaire. Feedback incorporation was analyzed by comparing the first and second essay versions and reviewing participants' interactions with ChatGPT, while pre- and post-task questionnaires provided insight into changes in participants' perceptions. Three distinct profiles of feedback incorporation emerged: Group A ($N = 2$) was characterized by integral incorporation - participants replaced their original drafts entirely by ChatGPT-generated text; Group B ($N = 3$) employed writing assistance strategies and selectively adopted suggestions to refine their writing; and Group C ($N = 2$) employed a partial incorporation, integrating specific excerpts into their original texts. Perceptions of ChatGPT varied across groups and appeared to be influenced by prior experience, the training session, and internal factors such as emotional reactions and social desirability bias. While Group A's stated beliefs remained unchanged despite actions that contradicted them, Group B demonstrated increasing interest and an apparent critical engagement with the tool, and Group C reported diminished trust in the tool due to issues with coherence and accuracy. These findings highlight the importance of fostering comprehensive AI literacy at all levels of education, emphasizing not only technical skills but also ethical awareness, critical evaluation of AI-generated content, and the risks of misinformation. The study reinforces the shared responsibility among educators, institutions, and policymakers in preparing students for critical and responsible engagement with AI tools, and it encourages further interdisciplinary research on human-AI interaction in language learning contexts.

Keywords: ChatGPT; automated writing evaluation; second language learning; AI literacy; human-AI interaction.

RESUMO

Este estudo exploratório investiga como estudantes de graduação brasileiros interagem com o feedback do ChatGPT e o incorporam durante uma tarefa de produção textual em inglês como segunda língua (L2), além de examinar como suas percepções sobre a ferramenta evoluem antes e depois da tarefa de feedback. Guiado por duas perguntas de pesquisa, o estudo analisa (1) como estudantes integram o feedback do ChatGPT em suas produções escritas em L2 e (2) como suas percepções sobre o ChatGPT como uma ferramenta de Avaliação Automatizada da Escrita (AWE, na sigla em inglês) mudam após a interação com o sistema. Adotando um desenho quase-experimental, transversal, e com pós-teste apenas, o estudo contou com sete estudantes de graduação (N = 7) de diferentes cursos de graduação da Universidade Federal de Santa Catarina (UFSC), todos com proficiência intermediária em inglês (nível B1-B2 do CEFR). A coleta de dados consistiu em uma redação manuscrita inicial em sala de aula, um questionário pré-tarefa, uma sessão de treinamento de uma hora e meia sobre o uso do ChatGPT (abordando funcionalidades, estratégias de formulação de prompts e considerações éticas), uma tarefa de revisão utilizando o feedback do ChatGPT e um questionário pós-tarefa. A incorporação do feedback foi analisada por meio da comparação entre as duas versões das redações e da revisão das interações dos participantes com o ChatGPT, enquanto os questionários pré- e pós-tarefa forneceram insights sobre mudanças nas percepções dos participantes. Três perfis distintos de incorporação de feedback emergiram: o Grupo A (N = 2) caracterizou-se pela incorporação integral, substituindo completamente seus rascunhos originais por textos gerados pelo ChatGPT; o Grupo B (N = 3) utilizou estratégias de assistência à escrita, adotando sugestões de forma seletiva para refinar seus textos; e o Grupo C (N = 2) fez uma incorporação parcial, integrando trechos específicos ao texto original. As percepções sobre o ChatGPT variaram entre os grupos e pareceram ser influenciadas por experiências prévias, pela sessão de treinamento e por fatores internos como reações emocionais e viés de desejabilidade social. Enquanto os membros do Grupo A mantiveram suas crenças inalteradas, mesmo com ações que as contradiziam, o Grupo B demonstrou crescente interesse e um aparente engajamento crítico com a ferramenta, e o Grupo C relatou perda de confiança na ferramenta devido a problemas de coerência e precisão. Esses dados destacam a importância de promover o letramento em IA em todos os níveis educacionais, enfatizando não apenas habilidades técnicas, mas também a consciência ética, a avaliação crítica de conteúdos gerados por IA e os riscos da desinformação. O estudo reforça a responsabilidade compartilhada entre educadores, instituições e formuladores de políticas na preparação dos estudantes para um engajamento crítico e responsável com ferramentas de IA, e incentiva novas pesquisas interdisciplinares sobre a interação humano-IA em contextos de aprendizagem de línguas.

Palavras-chave: ChatGPT; avaliação automatizada de escrita; aprendizagem de segunda língua; letramento em IA; interação humano-IA.

LIST OF FIGURES

Figure 1 – Grammarly advertisement showcasing its AWE.....	21
Figure 2 – A screenshot of a prompt submitted to ChatGPT and its response.....	27
Figure 3 – Three developmental stages of students’ use of ChatGPT in a writing assignment.	28
Figure 4 – Slideshow used during the training on incorporating LLM feedback on writing	35
Figure 5 – Study design flowchart	36

LIST OF TABLES

Table 1 – Groups divided by number of participants and the type of feedback incorporation.....	48
Table 2 – Excerpts demonstrating integral incorporation by Group A	50
Table 3 – Excerpts demonstrating Group B’s interactions.....	51
Table 4 – Excerpts demonstrating Group C partial incorporation	54

LIST OF ABBREVIATIONS AND ACRONYMS

ABNT	Associação Brasileira de Normas Técnicas
ADHD	Attention Deficit and Hyperactivity Disorder
AI	Artificial Intelligence
AIED	Artificial Intelligence in Education
AWCF	Automated Written Corrective Feedback
AWE	Automated Writing Evaluation
CAAE	Certificado de Apresentação de Apreciação Ética
CCE	Centro de Comunicação e Expressão
CEFR	Common European Framework of Reference
CEPSH	Comitê de Ética em Pesquisa com Seres Humanos
D&D	Dungeons and Dragons
FL	Feedback Literacy
GenAI	Generative Artificial Intelligence
GPT	Generative Pre-trained Transformer
HEI	Higher Education Institution
LLM	Large Language Model
L2	Second Language
LLE	Foreign Language and Literature Department code
ML	Machine Learning
MtG	Magic: The Gathering
NLP	Natural Language Processing
PPGI	Graduate Program in English: Linguistics and Literary Studies
RPG	Role Playing Game

TCLE	Termo de Consentimento Livre e Esclarecido
UFSC	Universidade Federal de Santa Catarina
UNESCO	United Nations Educational, Scientific, and Cultural Organization
WotC	Wizards of the Coast

TABLE OF CONTENTS

1. INTRODUCTION	14
1.1 OBJECTIVES	16
1.1.1 Main objective.....	16
1.1.2 Specific objectives	16
1.2 RESEARCH QUESTIONS	16
2. REVIEW OF LITERATURE.....	17
2.1 FEEDBACK AS A TOOL FOR LEARNING	17
2.2 FEEDBACK IN ACADEMIC WRITING	19
2.3 ARTIFICIAL INTELLIGENCE IN (LANGUAGE) EDUCATION: UNITED NATIONS POLICIES FOR AIED	21
2.4 AUTOMATED WRITING EVALUATION.....	25
2.5 CHATGPT AS AN AUTOMATED WRITING EVALUATION SYSTEM	26
2.6 LITERACIES, LANGUAGE EDUCATION, FEEDBACK, AND GENERATIVE AI	29
3. METHOD	32
3.1 PARTICIPANTS AND CONTEXT	32
3.2 STUDY DESIGN	33
3.3 DESCRIPTION OF THE SYSTEM AND ITS USE IN THE STUDY	36
3.3.1 ChatGPT for Writing Assessment.....	36
3.4 INSTRUMENTS AND PROCEDURES FOR DATA COLLECTION.....	37
3.4.1 Pre-task questionnaire.....	37
3.4.2 Text submission and ChatGPT interactions	38
3.4.3 Post-task questionnaire.....	39
3.5 DATA ANALYSIS	40
3.5.1 Participants' Feedback Incorporation	40
3.5.2 Participants' Perceptions of ChatGPT as an AWE	41
4. RESULTS AND DISCUSSION.....	43

4.1 PARTICIPANTS' INTERACTIONS WITH CHATGPT	43
4.3 PRE-TASK QUESTIONNAIRE	56
4.3.1 Group A - Integral incorporation	56
4.3.2 Group B - Writing assistance	57
4.3.3 Group C - Partial incorporation	59
4.4 POST-TASK QUESTIONNAIRE	60
4.4.1 Group A - Integral incorporation	60
4.4.2 Group B - Writing assistance	61
4.4.3 Group C - Partial incorporation	63
4.5 GROUPS ANALYSIS	64
4.5.1 Group A analysis	65
4.5.2 Group B analysis	67
4.5.3 Group C analysis	69
5. CONCLUSION	75
5.1 PEDAGOGICAL IMPLICATIONS	78
5.2 LIMITATIONS OF THE STUDY	79
5.3 SUGGESTIONS FOR FURTHER RESEARCH	81
REFERENCES	84
APPENDIX A - Termo de Consentimento Livre e Esclarecido (TCLE)	92
APPENDIX B - Pre-task writing test	97
APPENDIX C - Pre-task questionnaire	98
APPENDIX D - Training session class plan	105
APPENDIX E - Instructions for task development	107
APPENDIX F - Post-task questionnaire	108
APPENDIX G - Data sheets	116

1. INTRODUCTION

Writing constitutes a social practice: in most cases, texts are not produced in sterile, controlled environments under the supervision of an instructor. Instead, they are part of a broader message, which also conveys the purpose of the text, the context in which it was produced, and its intended audience, for instance (Hyland, 2007). However, when Composition Studies were formalized as an area of research in the United States in the 1950s, its focus was dominated by a prescriptivist perspective influenced by Formalism (Nystrand; Greene; Wiemelt, 1993).

In 1981, Flower and Hayes (1981, p.366) published the seminal paper *Cognitive Process Theory of Writing*, proposing a theory that approaches writing “as a set of distinctive thinking processes which writers orchestrate or organize during the act of composing”. Feedback is a central piece of the process approach to writing, as it generates input to improve the text (Keh, 1990). This feedback can be explicit or implicit and generated by an outer agent (such as a teacher, a peer, or a machine) or self-evaluation (Carless; Boud, 2018; Boud; Lawson; Thompson, 2013).

Writing instruction and feedback are central aspects of literacy development. The traditional approach to literacy reduced the concept to a mental and cognitive skill of interpreting and producing written language (Gee, 2015). In opposition to this perception, the New London Group (1996) proposed a new approach to language pedagogy due to the changing social environment of that period. Their article, *A Pedagogy of Multiliteracies: Designing Social Futures*, introduced the concept of multiliteracies, proposing that aspects such as the multiplicity of mediatic channels and the saliency of diversity in language and culture would require a broader and more comprehensible understanding of literacy studies (The New London Group, 1996).

Due to the global economic shift from Fordist mass production techniques of repetitive work towards post-Fordist flattened hierarchies, unskilled workers were replaced by multi-skilled ones (The New London Group, 1996). As command-and-control structures gave the floor to pedagogy, the multiliteracies perspective understood that the traditional literacy concept should be broadened to embrace new modalities of language production, such as music, gestures, and pictures (Sang, 2017). For the New London Group (1996, p.6), the reconceptualization of literacy studies was necessary because “as educators, we have a greater responsibility to consider the implications of what we do in relation to a productive working life”.

The multiplicity of aspects included in literacy studies was further broadened with the new characteristics of productive working life brought by the 21st century. Advancements in technology have further shaped academic literacies, including digital literacy (Zhang; Hyland, 2023), and reshaped the contexts in which feedback literacy is applied (Carless; Boud, 2018), both of which are explored further in the Review of Literature. Notably, Automated Writing Evaluation¹ (AWE) tools have gained prominence, offering “more accurate yet flexible affordances for automatic corrective feedback” (Barrot, 2021, p.584). Therefore, computer-generated Automated Feedback (AF), which once was portrayed as unreliable (Hyland; Hyland, 2006), is now related to a series of positive aspects of L2² writing development, such as the decrease of errors in longitudinal writing performance and the improvement of learners’ internal factors, such as motivation and self-confidence, for instance (Fan; Ma, 2022).

However, recent Artificial Intelligence (AI) developments, exemplified by ChatGPT (OpenAI, 2023), have changed the writing and instruction landscape. ChatGPT, a conversational chatbot that produces human-like text, is apt to deal with language-related assignments and has educational implications (OpenAI, 2023d; Kasneci *et al.*, 2023). The impacts of ChatGPT on Education have reverberated worldwide, and the United Nations Educational, Scientific and Cultural Organization (UNESCO) has discussed its developments, limitations, and orientations in guides and other documents (UNESCO, 2023a, 2023b), stressing the relevance of fostering research on this topic.

As ChatGPT gains prominence, research indicates that it may provide new ways of learning and teaching, including in L2 instructional contexts (Yan, 2023; Schlosser, 2023). However, it has called attention to the need for critical use of the tool due to its lack of regulation, cognitive bias, and discrimination (UNESCO, 2023). Therefore, this research aims to assess Brazilian undergraduate students’ adoption of ChatGPT for AWE in a writing assignment. By exploring students’ feedback incorporation and their reported experiences and perceptions while using AI for L2

¹ The terms Automated Writing Evaluation (AWE) and Automated Written Corrective Feedback (AWCF) are used interchangeably (Koltovskaia, 2020). However, in this text I use the term AWE, as I consider that Generative Artificial Intelligence tools (which is the scope of this thesis) may not necessarily provide a corrective feedback, rather than a writing evaluation, depending on the prompt provided by the user.

² In this thesis, I adopt the term Second Language (L2) to align with the literature that informs this study and to avoid the limiting connotations of foreign language as a language of “others” (Zaccaron, 2018).

learning purposes, this thesis seeks to contribute to the ongoing dialogue on using AI technologies in education.

1.1 OBJECTIVES

1.1.1 Main objective

To investigate Brazilian undergraduate students' interaction and incorporation of ChatGPT's feedback in an English as a Second Language (L2) writing assignment and their perceptions regarding its use before and after the feedback task.

1.1.2 Specific objectives

- To analyze the extent to which Brazilian undergraduate students incorporate ChatGPT's feedback into their L2 writing assignments, categorizing it according to incorporation patterns.
- To investigate students' perceptions and behavior toward ChatGPT use in L2 writing assignments before and after its use as a source of feedback in a written task.

1.2 RESEARCH QUESTIONS

Given the challenges involved in L2 academic writing and the relevance of multiliteracies, Automated Feedback, Automated Writing Evaluation, and ChatGPT in Higher Education, this exploratory study proposes two research questions:

RQ01 - How do Brazilian undergraduate students incorporate ChatGPT's feedback into an L2 writing assignment?

RQ02 - How do Brazilian undergraduate students' perceptions regarding ChatGPT as an AWE tool in an L2 assignment change after they engage with the feedback task?

2. REVIEW OF LITERATURE

2.1 FEEDBACK AS A TOOL FOR LEARNING

The basis for adopting feedback as a tool for learning in Second Language Acquisition (SLA) is the “robust connection between interaction and learning” (Gass; Mackey, 2015, p.181). The interaction approach combines aspects of the input hypothesis (Krashen, 1985) and the output hypothesis (Swain, 1985) to propose that negotiation for meaning is a relevant moment in L2 development because learners may notice the gaps in their interlanguage (Schmidt; Frota, 1986). While being exposed to peer feedback, learners interact in L2 and negotiate meaning. Thus, the interaction of speakers makes “problematic aspects of learners’ interlanguage salient and may give them additional opportunities to focus on their production or comprehension, thus promoting L2 development” (Gass; Mackey, 2015, p.187). This may occur via explicit feedback, consisting of corrections and metalinguistic explanations, or via implicit feedback, which includes strategies such as confirmation and comprehension checks, clarification requests, and recasts (Gass; Mackey, 2015).

Truscott (1996) challenged the notion that error correction aids L2 development. According to him, correcting grammar is ineffective due to the nature of correction and learning, and has harmful effects on L2 learners’ affective aspects. The author points out that the foundation for this issue is the “inherent unpleasantness of correction” (Truscott, 1996, p.355), which hinders the development of a positive attitude regarding language learning. In response, Ferris (1999) claimed that Truscott’s thesis was premature for different reasons. For the author, Truscott’s (1996) premises were not accurate, as his paper did not make a clear distinction between poor and effective error correction. Ferris (1999) also claimed that Truscott exaggerates negative evidence to support his thesis by comparing studies that involved non-comparable groups of subjects and differing research paradigms and teaching strategies.

In response to Ferris's (1999) critique, Truscott (1999) rebutted the criticism of the differentiation between poor and effective error correction. The rebuttal emphasizes that the claim aims to argue that students benefit from correction when it is clear and consistent, rather than whether some varieties of corrections have more problems than others. Truscott (1999) argues that, when Ferris (1999) pointed to the lack of evidence from the literature on Truscott's (1996) argument, Ferris did not indicate how their sources supported meaningful evidence on the benefits of error correction. Finally, although agreeing with Ferris that many questions regarding error correction remain open, Truscott (1999, p.121) defends that "correction might become an appropriate, if strictly limited, tool. It should be remembered, though, that the presence of beneficial effects is not, in itself, sufficient justification for correction".

Despite Truscott's (1996; 1999) critiques of corrective feedback (CF), studies conducted in the decades after their publication did not find detrimental effects of error correction on writing. A review of 15 empirical studies identified that "error correction is effective for short-term revision, and increasing evidence also suggests that both focused and unfocused CF can facilitate the acquisition of either a single or a wider range of grammatical features" (Meng 2013, p.81).

The paradigm towards language correction has shifted as feedback started to be considered a powerful tool for learning (Lipnevich; Panadero, 2021). Thus, it is now known that different types of feedback can be more or less effective depending on learners' perceptions and teachers' approaches. According to Hattie and Timperley (2007, p.104),

To be effective, feedback needs to be clear, purposeful, meaningful, and compatible with students' prior knowledge, and to provide logical connections. It also needs to prompt active information processing on the part of learners, have low task complexity, relate to specific and clear goals, and provide little threat to the person at the self level.

Lipnevich and Panadero (2021) reviewed 14 prominent feedback models and theories. The authors demonstrated that, regardless of the proliferation of models in this field of research, there is a common agreement that feedback is essential for improving student performance. They argue that research should describe how those models and theories behave in authentic classroom contexts via replicable studies in instructional settings.

In this section, I have examined how the field of research on feedback in L2 has moved from the debate on error correction (Truscott, 1996; Ferris, 1999; Truscott, 1999) to the multiple approaches, models, and theories of feedback. The following section discusses feedback in the context of academic writing, connecting it to a feedback-focused understanding of literacy that considers the multimodal and digital dimensions of language learning (Carless; Boud, 2018; The New London Group, 1996; Zhang; Hyland, 2022).

2.2 FEEDBACK IN ACADEMIC WRITING

Studies on corrective feedback for L2 writing instruction have gained prominence in the area of Academic Literacies (Hyland; Hyland, 2006; Yu; Lee, 2016). For Lea and Street (2006, p.158), “Academic literacy practices — reading and writing within disciplines — constitute central processes through which students learn new subjects and develop their knowledge about new areas of study”. Literacy, in a general sense, is embedded in academic, social, and cultural practices once readers and writers are primarily engaged in linguistic activities. Therefore,

written language never sits all by itself and it is rarely if ever fully cut off from oral language and action. Rather, within different practices, it is integrated with different ways of (1) using oral language; (2) of acting and interacting; (3) of knowing, valuing, and believing; and, too, often (4) of using various sorts of tools and technologies (Gee, 2015, p.36).

Within the scope of literacy studies, Carless and Boud (2018) discuss the relevance of feedback in literacy development. The authors define feedback as “a process through which learners make sense of information from various sources and use it to enhance their work or learning strategies” (2018, p.1315). The ability to critically understand, give, receive, and act upon feedback in different contexts (in academic or working contexts, for instance) and modalities (written or oral) is defined as Feedback Literacy (Sutton, 2012; Nieminen; Carless, 2023).

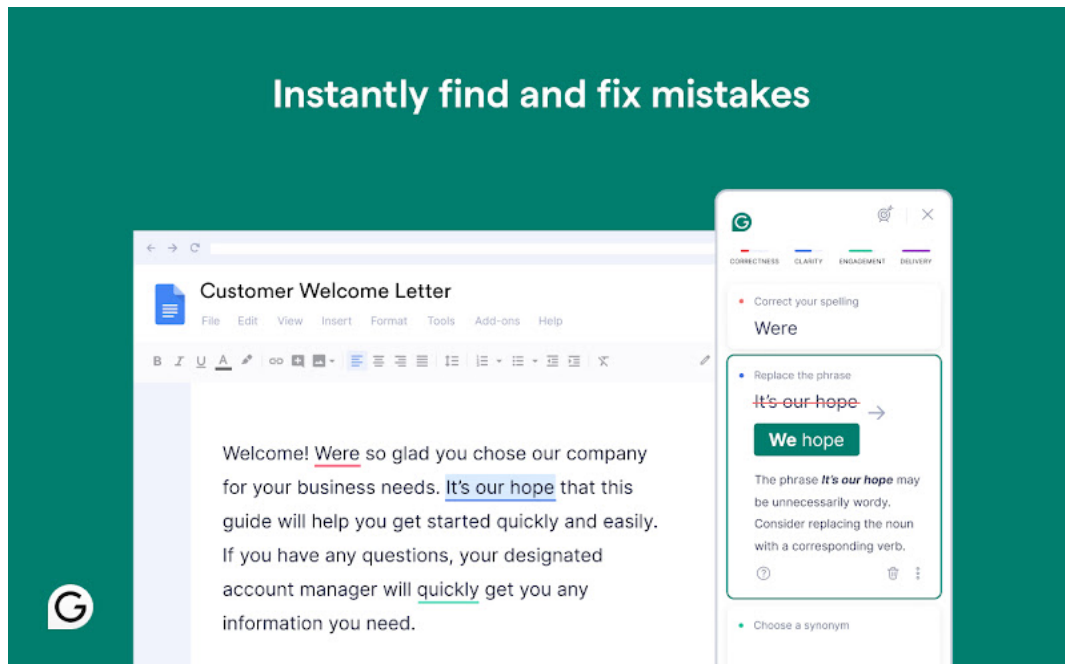
Research on feedback studies has explored the impact of different types of feedback on students’ L2 writing. Hyland and Hyland (2006) discussed the strengths and weaknesses of various kinds of feedback, one of them being Automated Feedback (AF), which consists of written feedback generated automatically by a machine or

system. The technology used in AF functioning mainly consists of statistical models that count the probability of association between two words and the likelihood that elements (words) in a sequence are independent. Using both measures, systems such as Criterion could examine whether words were associated and the “strength” of their association (Burstein; Chodorow; Leacock, 2004).

In this context, technology has played an important and progressive role in feedback processes and AF, as “e-feedback” (i.e., electronic feedback) tools became available to the public in general (Chang *et al.*, 2018). This caused feedback studies to move their primary focus from the effectiveness of error correction to investigate a new range of research topics, such as contextual constraints, individual, social, and cultural factors, and long-term impacts on students’ writing (Hyland; Hyland, 2006).

Artificial Intelligence (AI) technologies have fostered AF and e-feedback in the last decade, allowing Automated Writing Evaluation (AWE) tools to assess, correct, and improve texts (Fleckenstein; Liebenow; Meyer, 2023). Grammarly (Figure 1) is one example of an AI-enabled writing assistant that works across different platforms, devices, applications, and websites. It instantly generates Automated Written Corrective Feedback by underlining imprecise or inadequate excerpts. Thus, it promises to provide “feedback on spelling, punctuation, grammar, and conventions, including spacing, capitalization, and dialect-specific spelling” (Koltovskaia, 2020, p.4). Therefore, considering the impact of AI on AWE systems development, the next subtopic aims to define AI and its primary constructs and connect it to the development of feedback literacy (FL).

Figure 1 – Grammarly advertisement showcasing its AWE.



Source: Grammarly, 2024.

2.3 ARTIFICIAL INTELLIGENCE IN (LANGUAGE) EDUCATION: UNITED NATIONS POLICIES FOR AIED

First, to comprehend Artificial Intelligence in Education (AIED) and its impacts, it is necessary to present some terms that underlie this concept. Foremost, Artificial Intelligence (AI) is considered “[...] a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions.” (Council of Europe, 2024, p.7). Machine Learning (ML), in turn, is one type of AI able to identify patterns and draw inferences based on large data sets, often referred to as Big Data (Holmes *et al.*, 2023). Finally, Natural Language Processing (NLP) systems use ML and Big Data “to transcribe, interpret, translate and create text and spoken language” (Holmes *et al.*, 2023, p.7).

The diverse applications of Artificial Intelligence in Education may benefit a range of educational processes. For example, AI tools can provide individualized feedback and adaptive instruction, besides supporting students with special educational needs, identifying patterns in academic progress, and helping to decrease part of teachers' workload with assessments, for instance (UNESCO, 2023b).

Nonetheless, the United Nations Educational, Scientific and Cultural Organization (UNESCO) argues that students, teachers, and institutions should be cognizant of the emergence of AI literacy, which is the understanding of AI's "basic functioning, potential benefits, and challenges" (UNESCO, 2023a, p.48).

Therefore, the development of AI literacy emerges as an essential part of incorporating AI into Education. Talking specifically about the role of Higher Education Institutions (HEIs), those institutions will be responsible for a new mission of educating the general public about AI once the skills related to the fair and ethical employment of AI in real-world scenarios need to be developed to build an AI literate society (UNESCO, 2023b).

UNESCO has been responsible for laying the foundation stone of international policies on the use of AI in Higher Education. The starting point of the process of policy development was the International Conference on Artificial Intelligence and Education, held in 2019. The conference's outcome document was the *Beijing Consensus on Artificial Intelligence and Education*, adopted by Member States and partner organizations (UNESCO, 2019a). This document represented the first worldwide agreement on educational policies for AI that offered guidance to policymakers, researchers, teachers, and stakeholders.

The Beijing Consensus emphasized the urge for United Nations Member States to define policies and strategies regarding the use of AI in their educational systems. When determining the scope of AI competencies for teachers and students, policymakers should consider the "systemic and long-term transformation of the labor market, including its gender dynamic, due to AI adoption" (UNESCO, 2019b, p. 6). Thus, this document indicates that AI should be employed to ensure the improvement of education worldwide, and international organizations should coordinate actions to promote an equitable development of AI technologies.

In 2021, UNESCO published *AI and education: guidance for policy-makers*, a guide instructing policymakers on best practices to implement policies for AI in education. Amongst the document's proposals, it recommends "to take institutional actions to enhance AI literacy across all sectors of society" (UNESCO, 2021, p. 36), guaranteeing that citizens should be educated to be critical and responsible for their choices and cognizant of the effect of AI on their lives. Moreover, the document orients pilot testing on AI initiatives, aiming to build evidence-based knowledge and foster

transferable skills such as problem-solving and creativity, for instance (UNESCO, 2021).

In the same year, UNESCO published the *Recommendation on the Ethics of Artificial Intelligence*, which addresses the need for ethical recommendations to use AI for the good of humanity and the environment. On the recommendation for educational policies, the document stresses that

Member States should promote the acquisition of “prerequisite skills” for AI education, such as basic literacy, numeracy, coding and digital skills, and media and information literacy, as well as critical and creative thinking, teamwork, communication, socio-emotional and AI ethics skills, especially in countries and in regions or areas within countries where there are notable gaps in the education of these skills (UNESCO, 2021, p.33).

This excerpt reinforces UNESCO’s concern about integrating traditional skills (as the “prerequisite skills” mentioned above) into AI literacy, recognizing the importance of human interaction within such technologies. Besides that, monitoring and evaluation are also continuous aspects of all policy documents mentioned until this point, which highlights that member states should monitor and evaluate AI policies via empirical research, empirical repositories, and observatories.

After the release of ChatGPT in 2021, UNESCO published a *Quick Start Guide on Using ChatGPT and Artificial Intelligence in Higher Education*. Among its leading suggestions, the document emphasizes the relevance of introducing “clear guidance for students and instructors about how and when ChatGPT can be used (and when it cannot)” (UNESCO, 2023a, p.13). Besides that, it reinforces the idea of training teachers, researchers, and students to cope with AI, its potential, and limitations for teaching, learning, and researching. As it resonates with other UNESCO documents (2023b, 2021, 2019a), Core AI competencies and skills should be developed to guarantee academic integrity and the development of transferable skills.

The landscape of AI skills in the labor market is detailed in *Harnessing the Era of Artificial Intelligence in Higher Education: A Primer for Higher Education Stakeholders*. UNESCO (2023b, p.44) argues that “some tasks are taken over by algorithms, and new tasks are being created, in part, to create and manage these algorithms”. However, it is interesting that it is expected that AI also leads to the creation of jobs not related to AI itself, but in areas not mastered by AI. These roles

are related to creativity, reasoning, critical thinking³, and social and interpersonal skills — tasks that artificial agents cannot fully accomplish.

Aiming to specify the key AI competencies that should be developed by students, UNESCO published a framework for the development of AI skills. In the *AI competency framework for students*, UNESCO (2024, p.27) suggests that “students’ AI competencies are the result of the integrated interventions of national AI curricula; extracurricular programs; informal learning through various media, including the internet; and engagement with families and local communities.”

Thus, they propose a three-level approach: 1 - at the “Understand” level, students should understand what AI is and develop age-appropriate interpretations of the technology; 2 - at the “Apply” level, students should develop solid and transferable conceptual knowledge about AI based on a human-centered point of view; and 3 - at the “Create” level, students should acquire advanced competencies to create, configure and deploy AI solutions considering social responsibilities regarding AI.

Therefore, UNESCO has advocated for policies related to ethical, human-centered, and socially responsible use of AI in education. The central theme of the policy documents was the development of AI skills among students, teachers, researchers, and HEIs’ staff. Besides that, UNESCO calls attention to developing skills that AI does not excel in and that are uniquely human, such as critical thinking, for instance. These attributes tend to be valued as machines will not be able to replace them, and HEIs should be attentive to the core competencies needed to face the challenge of the economic and labor market shift to AI-driven economies.

Among the technologies associated with writing development and fostered by AI technologies, automated writing evaluation (AWE) tools have undergone a significant modernization in the last decades, providing instant and adaptable Automated Feedback (AF) on online word processors. The next subtopic will approach automated writing tools based on AI, analyzing empirical studies on the use of such tools in pedagogical contexts and how they may impact learners’ writing development.

³ In this thesis, critical thinking is understood as “perceiving the root causes of one’s place in society—the socioeconomic, political, cultural, and historical context of our personal lives” (Auerbach; Wallerstein, 1987, p.1).

2.4 AUTOMATED WRITING EVALUATION

Automated Writing Evaluation refers to computer-generated feedback provided to a piece of writing for instructional or assessment purposes (Barrot, 2023). These tools are based on AI, ML, and NLP, and can evaluate, interpret, and generate comments on students' writings (Fleckenstein; Liebenow; Meyer, 2023). In a systematic review, Fan and Ma (2022) reviewed 22 studies (carried out between 2005 and April 2020) regarding the effects of AWE in writing development. Their findings demonstrated that AWE tools yielded positive outcomes, as evidenced by the decrease in errors in revised and new works, and improvements in writing skills in long-term performance. Moreover, the authors identified other affective and emotional aspects, such as an increase in motivation, self-confidence, and enjoyment, and a decrease in writing apprehension, for instance (Fan; Ma, 2022).

Fleckenstein, Liebenow, and Meyer (2023) proposed a multi-level meta-analysis examining the 84 effect sizes of 20 studies investigating AWE in secondary and tertiary education contexts. They found a medium effect size, supporting AWE feedback for students' L1 and L2 writing enhancement. The effect size of AWE on L2 writers was large, and for L1 writers, it was medium. Therefore, the authors indicate that "the heterogeneity in the results suggests that automated feedback may not have the same impact on all students" (Fleckenstein; Liebenow; Meyer, 2023, p.09), which supports the idea that different levels of literacy development may play a role in feedback uptake.

Moreover, other studies have explored the use of AWE in academic contexts. Zhang and Hyland (2023) investigated the influence of digital literacy on the ability to employ AWE tools' feedback. According to the authors, digital literacy "refers to an individual's ability to evaluate and compose information appropriately through writing and other media using digital platforms" (Zhang; Hyland, 2023, p.3). The authors found out that the effective employment of AWE feedback technological tools by L2 students is related to their level of digital literacy, and this "involves behavioural, affective, and cognitive engagement, which is closely connected to 'thinking', 'meaning', and 'doing' aspects of digital literacy" (Zhang; Hyland, 2023, p.20).

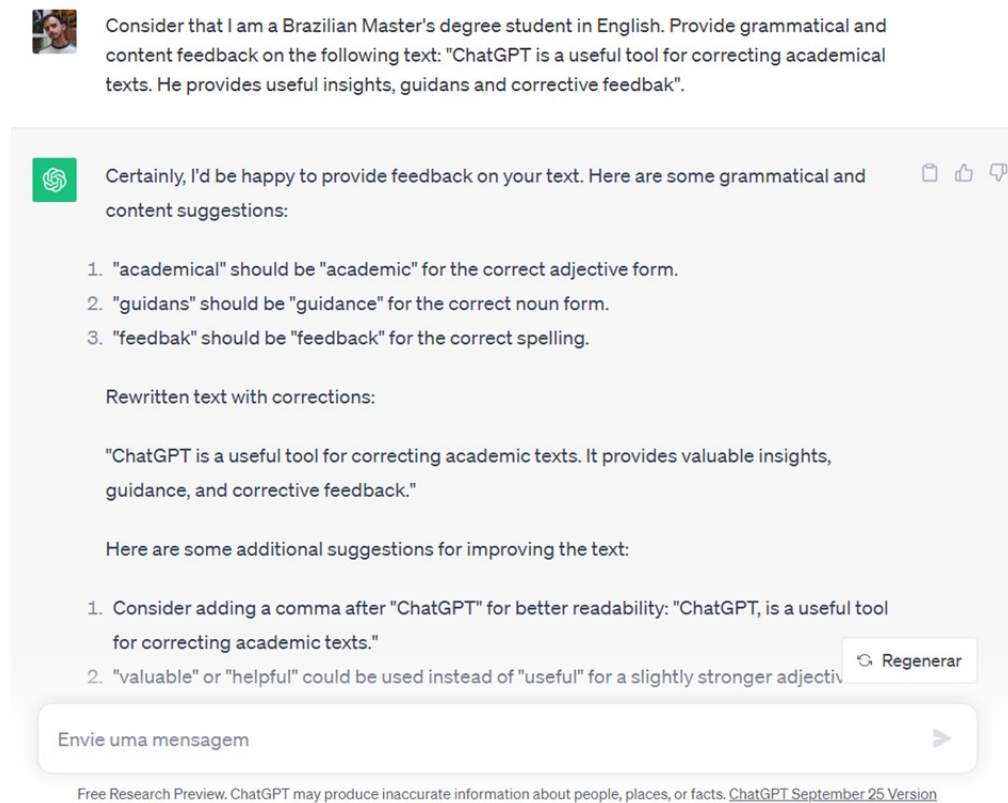
Studies considering different AWE tools, such as Criterion (Link; Mehrzad; Rahimi, 2023) and Grammarly (Barrot, 2021), have demonstrated positive outcomes

in providing instant and accurate feedback, besides improving the users' internal factors, such as engagement and motivation. More recently, studies have investigated the use of Generative Artificial Intelligence (GenAI) powered chatbots for AWE (Yan, 2023; Su; Lin; Lai, 2023). The term GenAI "has been applied more precisely to models that may produce new, previously unseen information dependent on the data on which they were trained" (García-Peñalvo; Vázquez-Ingelmo, 2023, p.7), enabling chatbots to generate human-like text and convey it conversationally, allowing users to chat with the system (Cotton; Cotton; Shipway, 2023). The next subtopic details the use of a GenAI-based chatbot, ChatGPT, as a tool for automated written feedback in HEI contexts.

2.5 CHATGPT AS AN AUTOMATED WRITING EVALUATION SYSTEM

ChatGPT is a web-based system developed via the Generative Pretrained Transformer (GPT; Cotton; Cotton; Shipway, 2023), which is a Large Language Model (LLM). LLMs are language models "trained on massive amounts of text data and can generate human-like text, answer questions, and complete other language-related tasks with high accuracy" (Kasneci *et al.*, 2023, p.1). Figure 2 demonstrates how ChatGPT works when asked to provide feedback on grammar and content for a text that presents grammatical errors. In this example, the system offers a numbered list for each textual correction and complements its feedback with written advice for punctuation and vocabulary, for example. As a language model, it is important to stress that each time a query is submitted, ChatGPT will guess its response differently, whether the request is the same or not.

Figure 2 – A screenshot of a prompt submitted to ChatGPT and its response.



Source: OpenAI, 2023d.⁴

As it is possible to see in Figure 2, the system may provide inaccurate orientations, such as its first suggestion for text improvement: "Consider adding a comma after "ChatGPT" for better readability". In this example, the system suggested inserting a comma between the subject and the verb, which is incorrect, considering the standard written English. In terms of assessment, OpenAI (2023a) points out that

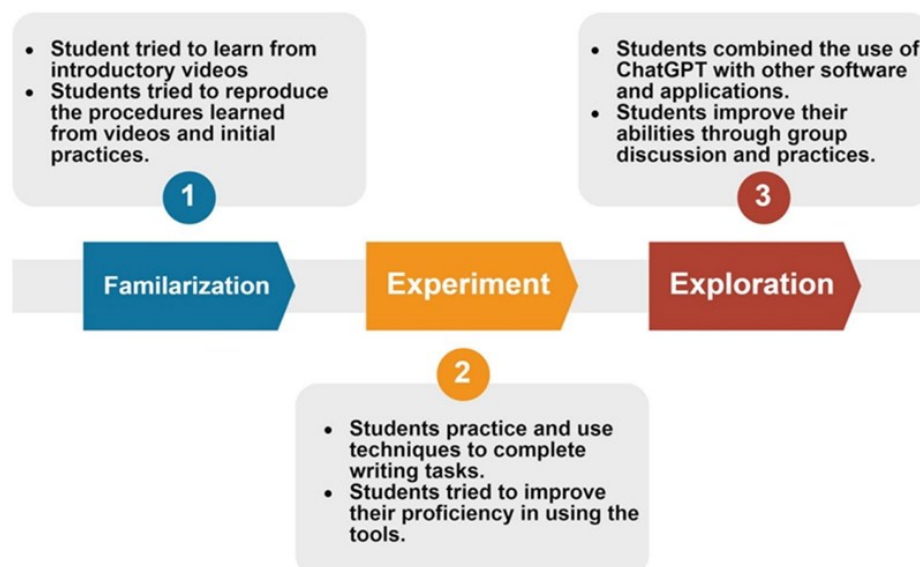
It is inadvisable and against our Usage Policies to rely on models for assessment decision purposes without a "human in the loop" (i.e., a person who may use AI as an aide, but who ultimately makes the decision using their own judgment). Models today are subject to biases and inaccuracies, and they are unable to capture the full complexity of a student or an educational context. Others have done research suggesting that ChatGPT can be a helpful tool, alongside teachers, for providing students with feedback. The model can provide another perspective on students' work and can help students think through and improve their drafts; teachers can use it to accelerate giving students robust and detailed feedback. For additional insight into how an entire class is doing, you can give ChatGPT all of a class's responses and ask it to identify common themes and common areas of strength and weaknesses.

⁴ This screenshot was made on ChatGPT 3.5, which finished training in early 2022 (OpenAI, 2022).

Likewise, empirical studies have reported ChatGPT use as an AWE tool in educational contexts. The main findings highlight the necessity of training teachers' and students' awareness of ChatGPT's strengths and weaknesses to maximize its benefits while measuring its pitfalls (Su; Lin; Lai, 2023). Similarly, Chinese doctoral students exposed to ChatGPT reported feeling the need for guidance for using ChatGPT regarding prompt writing, dealing with plagiarism risk, "and how to use ChatGPT as a self-learning tool for the concurrent development of writing and thinking" (Zou; Huang, 2023, p.15).

In this direction, Yan (2023) investigated the use of ChatGPT in a one-week ESL undergraduate practicum in China, developed specifically for introducing ChatGPT in ESL writing. The author indicates that students tended to go through three developmental stages: familiarization, experimenting, and exploration, detailed in Figure 3. In this sense, bearing in mind the pedagogical settings in which ChatGPT can be employed, Yan (2023, p.17) suggests that "Instead of obstructing students' exposure and access to such tools, lecturers should develop and implement a protocol of application to make proper use of the technology. Therefore, instructional guidance should be provided for using such technologies, and regulatory policies regarding plagiarism should be reinforced in HEIs".

Figure 3 – Three developmental stages of students' use of ChatGPT in a writing assignment.



Source: Yan, 2023, p.10.

Different studies described students' perceptions and use of AWE tools and ChatGPT for feedback. Schlosser (2023) identified that undergraduate students in first and second years in the Business, Marketing, and Advertising programs in Spain evaluated ChatGPT more positively as compared to the other two AWE systems (Grammarly and Write&Improve), and the "adjectives used to describe the feedback included "accurate, useful, detailed, structured and insightful" (Schlosser, 2023, p.28).

Basic *et al.* (2023) compared essay grades of 18 second-year Master's degree students in Croatia, divided into control and experimental groups. The experimental group used ChatGPT as a writing assistant, and their essays had an average lower score than the control group when evaluated by teachers, besides presenting more cases of plagiarism (Basic *et al.*, 2023). However, the authors did not specify how they prepared students for using the tool, which hinders the analysis of the influence of other factors, such as literacy, in general, on experimental group performance. Besides that, the limited number of participants (nine in each group) also stands out as a significant limitation of the research (Basic *et al.*, 2023).

Considering the lack of consensus in the literature regarding the use of ChatGPT as an AWE, I agree with UNESCO (2023b, p.28) that it is relevant "to question why students may turn to ChatGPT to produce assessments and what Higher Education Institutions (HEIs) can do to promote academic integrity and the importance of learning". Moreover, considering the limitations of this field of research (Yan, 2023; Schlosser, 2023; Basic *et al.*, 2023), the relevance of the present study is given by the necessity of understanding the effects of AI tools in literacy and writing development in a broad sense, including the paradigms now offered by GenAI, such as academic integrity and AI literacy, for instance (UNESCO, 2023a). The next subtopic aims to intertwine the theoretical underpinnings brought up until this point with the researcher's point of view on GenAI and its educational, social, and informational impacts.

2.6 LITERACIES, LANGUAGE EDUCATION, FEEDBACK, AND GENERATIVE AI

In this thesis, multiliteracies (The New London Group, 1996) embrace two main perspectives. One of these perspectives is Feedback Literacy (FL), which proposes that students should be able to receive, interpret, act upon, and provide feedback

(Carless; Boud, 2018). FL posits as a relevant aspect of Second Language Acquisition once it indicates that feedback may contribute to L2 learning, including writing development (Zaccaron; Xhafaj, 2020).

The second perspective highlights how the use of GenAI as a source of corrective writing feedback will be carried out. This thesis approaches it through the lens of AI literacy, which I conceptualize as the understanding of Generative AI principles and its basic functioning as a means to act critically and responsibly when navigating, living, learning, and working in the digital world through AI technologies (Ng *et al.*, 2021; UNESCO, 2023a). In addition, multiliteracies is embraced as a central concept in this perspective, because it provides a language development research framework that connects the multiplicity of communications channels and media (The New London Group, 1996), including emergent technologies such as GenAI.

Although the idea of (multi)literacies (The New London Group, 1996; Gee, 2015) was developed decades before the current possibilities offered by GenAI in the 21st century, it is still aligned with the discussions regarding the use of AI tools — such as Automated Writing Evaluation (AWE) tools — in learning/assessment⁵. As aforementioned, the effectiveness of these tools will depend on the level of literacy development and the social and cultural contexts in which they are used and taught, besides the individual factors, which will rely on aspects of students and teachers who employ them (Zhang; Hyland, 2023, on digital literacy, for instance).

However, it is essential to be critical when approaching AI and its risks, limitations, and benefits. Crawford (2023, p.37, our translation) defends that “the creation of contemporary AI systems relies on the exploration of the planet’s energetic and mineral resources, cheap workforce, and large scale data.” Therefore, it is unethical and unadvisable to research and write about AI, considering it an abstract and theoretical system. After all, AI relies upon and influences physical and tangible resources - its energetic consumption, effect on workers and economies, and its power to increase and/or decrease inequality worldwide.

Besides AI's environmental and social effects, understanding the ethical issues involved in GenAI systems conception and use is equally important. If students should

⁵ “The New Literacies Studies views different digital tools as technologies for giving and getting meaning, just like language [...] the meanings to which these technologies give rise are determined by the social, cultural, historical, and institutional practices of different groups of people. And, as with the NLS, these practices almost always involve more than just using a digital tool — they involve, as well, ways of acting, interacting, valuing, believing, and knowing.” (Gee, 2015, p. 44)

have their AI literacy skills trained, they must comprehend the bias embedded in AI systems. Bender *et al.* (2021, p.613) explain that the datasets used by GenAI to generate text and images have been shown to “encode stereotypical and derogatory associations along gender, race, ethnicity, and disability status.” According to the authors, dataset sizes do not guarantee they are diverse — that is, biases may be replicated in language models, as in mainstream media.

Underrepresented groups often disrupt hegemonic communications as they (re)shape digital narratives according to their points of view. Bender *et al.* (2021, p.614) mention how social movements produce new meanings and ways of communication in digital environments

For instance, the Black Lives Matter movement (BLM) influenced Wikipedia article generation and editing such that, as the BLM movement grew, articles covering shootings of Black people increased in coverage and were generated with reduced latency. Importantly, articles describing past shootings and incidents of police brutality were created and updated as articles for new events were created, reflecting how social movements make connections between events in time to form cohesive narratives.

Consequently, the dynamism of disputes around discourses poses challenges to language models, as their datasets may supply older and antiquated content regarding a specific stigmatized group. Bias may appear not due to the conception of GenAI systems itself, but rather due to the social prejudice and inequality of data that users produce online. Approaching AI literacy in language instruction involves more than knowing how to use GenAI systems; it also involves understanding why they operate the way they do. As Coeckelberg (2020, p. 53) points out, “The use and development of technology take place in a social context,” so teaching AI can not disregard the social aspects embedded in language production, especially in digital environments.

Considering the social, ethical, and pedagogical changes that AI may pose, this thesis investigates how undergraduate students employ, perceive, and act upon a GenAI tool in an L2 writing development context. Therefore, besides assessing students’ feedback incorporation, this work aims to understand how their beliefs, behavior, and previous knowledge on AI influence their use of ChatGPT. Bearing all of this in mind, the following section presents the study methodology.

3. METHOD

This exploratory qualitative study has a cross-sectional and quasi-experimental posttest-only design (Rogers; Révész, 2020) to examine how undergraduate students incorporate ChatGPT's feedback into their texts and what their perceptions of using ChatGPT as an AWE tool are. The independent variable manipulated is the use of ChatGPT as an AWE tool, which includes the training participants received on interacting with the tool and its subsequent feedback on their essays. The dependent variables are the incorporation (uptake) of ChatGPT's feedback, analyzed by the modifications participants made to the first draft of their essays based on the feedback received, and participants' perceptions of ChatGPT, assessed via questionnaires administered before and after the training.

This study was approved in the *Comitê de Ética em Pesquisa com Seres Humanos* (CEPSH/UFSC) under the process number 79678724.0.0000.0121 (CAAE number), following the national "guidelines and regulatory standards for research involving human beings" (Conselho Nacional de Saúde, 2012, p.1) indicated in the *Resolução CNS 466/12*. Before the beginning of the study, participants signed the consent form (*Termo de Consentimento Livre e Esclarecido*, Appendix A). This document indicates the free and informed consent of the participants who take part in the study, as well as the project objectives and other necessary information for its complete comprehension (Conselho Nacional de Saúde, 2012).

3.1 PARTICIPANTS AND CONTEXT

From the eight participants who joined the training session and interacted with ChatGPT, one did not submit the final version of their text. Consequently, seven participants who made modifications to their essays ($N = 7$) participated in the study. They were enrolled at Universidade Federal de Santa Catarina (UFSC) at the undergraduate level in the following programs: Foreign Languages (English and German), Law, Librarianship, Biology, and Executive Secretarial Sciences. The inclusion criteria were that students must have an intermediate level of proficiency in English (B1 level, according to the Common European Framework of Reference -

CEFR), as the tasks involved required sufficient L2 competence to produce essays in English and comprehend the feedback provided by ChatGPT.

The study was conducted within the context of the course LLE 7409 - Tópico Especial em Língua Inglesa II. This is a 72-hour course that has its program open, so that instructors can decide, depending on the term, what they wish to focus on. That means it may be offered in a specific semester, for example, with a focus on English as Lingua Franca, and in the next semester with a focus on Games for L2 learning. In the semester that data was collected, the focus was on academic writing in English. Along the course, the students (re)wrote essays, summaries, and a short literature review. Since a minimum B1 level of English was required for students to take part in the course, on the first day of classes, besides introducing the course, the teacher also applied an informal diagnostic test to evaluate whether all students were likely to be able to engage with the tasks demanded in the course. The group had 13 students enrolled, and 8 of them took the test (the others were absent from the first class). All these 8, and others whose data is not included in the study, agreed to contribute with data for this investigation, handed in the two versions of the essay, took part of the workshop of training for the use of ChatGPT, and did not miss more than one class of the course in the weeks between the day they handed in their first essay and the day they worked on their final version of it. According to the teacher's assessment of their performance in the diagnostic text and their first version of the essay, all eight students had a good level of English (3 were at the B1 level and five were closer to the B2 level).

3.2 STUDY DESIGN

The design for data collection was divided into six parts (Figure 4). First, participants produced an essay about a topic of their choice in English. To prevent participants from using any type of AWE tool or GenAI, the text was written on paper by hand and under supervision on the first day of classes, immediately after the diagnostic test. This occurred in a classroom in Centro de Comunicação e Expressão (CCE) at the Universidade Federal de Santa Catarina (UFSC - Campus Florianópolis). The class lasted four hours, with a 15-minute break between the two halves. Participants received brief instructions regarding text length and genre (Appendix B). Secondly, after delivering their drafts of the essay, participants answered a

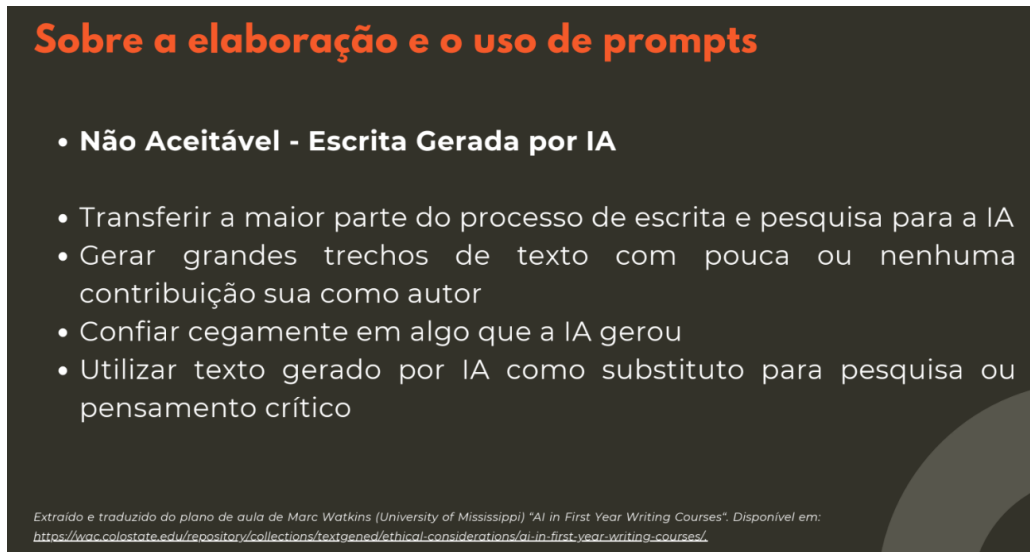
questionnaire in Portuguese, encompassing profile questions as well as their perceptions and previous knowledge of ChatGPT (Appendix C). This questionnaire was also filled in on paper and by hand, as it was administered in class right after the participants finished their first draft.

Five weeks later, participants engaged in the third phase of the study. They attended a one-and-a-half-hour training session on the use of ChatGPT feedback in academic writing, which comprised the “Familiarization” and “Experiment” steps of engaging with ChatGPT, as detailed by Yan (2023) in Figure 3 (see the training session plan in Appendix D). The training session occurred in a computer lab at CCE/UFSC and was conducted in Portuguese in an attempt to ensure that all students would comprehend the input from the training session, regardless of their level of L2 proficiency⁶. The training consisted of a slideshow presentation that brought information about AI, LLM, and, more specifically, ChatGPT, along with prompts that the researcher, the professor who was responsible for the course, and the participants used to discuss the acceptable uses and limitations of using ChatGPT for academic writing.

Although issues related to plagiarism had already been discussed earlier in the course, they were revisited during the training session, and participants were reminded of unacceptable practices when using ChatGPT for educational purposes (Figure 4). Finally, still as part of the training session, participants were invited to familiarize themselves with and test ChatGPT by asking the system to create a polemic tweet, with a maximum of 280 words, about a polemic or controversial topic of their choice, and then rewrite prompts for improving it. This warm-up activity aimed to prepare students to interact with the system as it tests the ethical implications of using LLMs.

⁶ The English proficiency level required to attend the course was B1 (CEFR). However, as L2 proficiency levels may vary among participants, the researcher decided to conduct the instructional session and the questionnaires in Portuguese. This condition aims to prevent the possible influence of participants with higher levels of L2 proficiency (a variable which is not being moderated) outperforming less proficient participants on input comprehension or output production.

Figure 4 – Slideshow used during the training on incorporating LLM feedback on writing



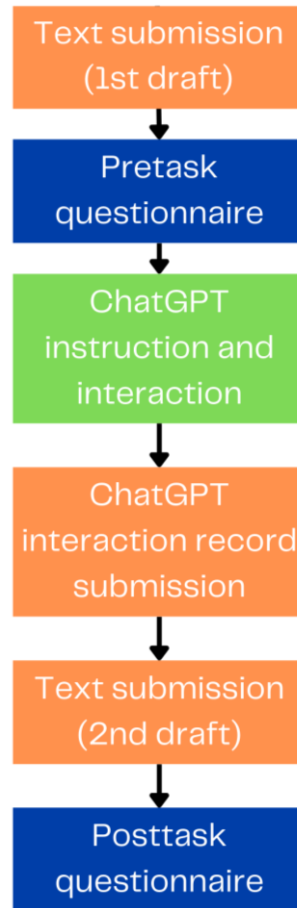
Source: the author, 2024.

The fourth part of the data collection occurred right after the training. In the computer lab, each participant had access to a computer connected to the internet for individual use. Then, participants accessed a computer and received a typed version of their first version of the essay (which they had written by hand in the first part of the study) in a text (.txt) file and worked on its improvement - a final version of the essay - using ChatGPT (check topic 3.4.2 for further details).

In the assessment step — the fifth part of the data collection —, participants received further instructions about using ChatGPT as a feedback provider for issues related to language use/linguistic accuracy and content (slideshow used for instructions in Appendix E). Students had 1 hour and 30 minutes to accomplish the task. Finally, on the sixth and last step, participants uploaded their final version of the essay, sent a link for the register of their interaction with ChatGPT, and answered a survey assessing their evaluation of using ChatGPT as an AWE (Appendix F).

After data collection, pre- and post-feedback texts and participants' interactions with ChatGPT were compared to identify modifications made by participants in their essays based on the feedback given by ChatGPT. Furthermore, comparing pre- and post-task answers to the surveys allowed the researcher to gain insights concerning changes in participants' reported perceptions and use of ChatGPT after they had a training session for its use in academic writing and used the feedback from the tool in the writing of the final version of their text, which answered RQ2.

Figure 5 – Study design flowchart



Source: the author, 2025.

3.3 DESCRIPTION OF THE SYSTEM AND ITS USE IN THE STUDY

3.3.1 ChatGPT for Writing Assessment

ChatGPT is an Artificial Intelligence-powered chatbot released in 2022 that can instantly generate human-like texts (OpenAI, 2023c). In this study, the system was used as an AWE tool, responding to prompts written by the students asking for feedback on their texts on form and content. OpenAI (2023b) points out that current LLMs “are subject to biases and inaccuracies, and they cannot capture the full complexity of a student or an educational context”. On the other hand, the company considers that “the model can provide another perspective on students’ work and can

help students think through and improve their drafts” (OpenAI, 2023b). Therefore, instruction on the system functionalities and limitations was provided to ensure that all participants interacted with core AI literacy skills (i.e., prompt writing and basic GenAI understanding) before completing the task.

The one-hour training on ChatGPT’s use was divided into two parts. In the first part, which lasted around 30 minutes, participants discussed ChatGPT’s basic functioning and some ethical issues regarding its use in Academia, mainly issues related to authorship. The materials for the instruction sessions were based on UNESCO (2023b; 2023a) documents named *Harnessing the Era of Artificial Intelligence in Higher Education: A Primer for Higher Education Stakeholders* and *ChatGPT and Artificial Intelligence in Higher Education: Quick Start Guide*. As presented in section 2.3, these two documents provide theoretical implications and examples of using ChatGPT and AI in HEIs. The second part, which lasted 30 minutes, focused on activities aiming to explore prompt writing, an essential step for completing the feedback task.

3.4 INSTRUMENTS AND PROCEDURES FOR DATA COLLECTION

3.4.1 Pre-task questionnaire

The pre-task questionnaire encompassed three sections. The first aimed at identification, and collected data about factual questions, such as age, the undergraduate program participants were associated with, and for how long they had been studying English. The second section surveyed participants with behavioral questions, and the third section approached attitudinal questions related to the use of AWE, AI, and ChatGPT. According to Dörnyiei (2011, p.102)

Behavioural questions [...] are used to find out what the respondents are doing or have done in the past, focusing on actions, lifestyles, habits, and personal history. Attitudinal questions [...] are used to find out what people think, covering attitudes, opinions, beliefs, interests, and values.

An example of a behavioral question from the second section is “Do you use any automated writing correction tool?”⁷. One of the attitudinal questions is a Likert-scale question: “Evaluate questions below from 1 to 6, being 1 (strongly disagree) and 6 (strongly agree): [...] It is acceptable to copy parts of the answers given by ChatGPT without indicating the source”⁸ — see Appendix C for the complete questionnaire.

3.4.2 Text submission and ChatGPT interactions

The first version of the essay was written on paper in a classroom at CCE/UFSC during the first week of the academic semester. The researcher transcribed participants’ texts to a Text File Format (.txt) in Windows Notepad. During ChatGPT’s interaction, participants need to use a computer with access to the internet to perform the task. Therefore, the text transcription to Notepad was an attempt to prevent participants from being influenced by automated text correction of more robust text processors such as Microsoft Word.

The second version of the text was produced by participants five weeks after the first draft, immediately after the feedback session. In the same computer lab where the training occurred, participants had one and a half hours to work on their second version using an individual computer in a computer lab at CCE/UFSC. They were instructed to access a link for a Google Drive folder containing their first drafts transcribed in a .txt format, as previously detailed. Participants were provided with login credentials to access ChatGPT and were advised by the researcher to improve their text using ChatGPT, based on the discussions and instructions received during the training. The ChatGPT version used was 3.5, which finished training in early 2022 (OpenAI, 2022).

After working on their texts, participants received a link to a Google Form. In this form, they pasted a link to their interaction with ChatGPT, generated via the “share

⁷ Original question: “6. Você usa alguma ferramenta de correção automática de texto? * (Ferramenta de correção do Word/Google Docs, Grammarly, dentre outros, por exemplo)” (Appendix C).

⁸ Original question: “Avalie as questões abaixo de 1 a 6, sendo 1 (discordo plenamente) e 6 (concordo plenamente): [...] 15. É aceitável copiar partes das respostas dadas pelo ChatGPT sem referenciar a fonte” (Appendix C).

chat” feature⁹. When clicking on the “share” icon on a ChatGPT interaction, a link was transferred to the user’s clipboard. Participants pasted the link generated together with their final text via a Google Docs questionnaire. Finally, participants submitted a copy of their interaction by copying and pasting it into the post-task questionnaire file to create a backup of their chat data, avoiding data loss. When completing the text submission, participants were redirected to the post-task questionnaire, detailed in the next section.

3.4.3 Post-task questionnaire

The post-task questionnaire (Appendix F) was also organized into three sections: in the first section, it collected the participants’ identification data, such as their name, the link of their interaction with ChatGPT, and the final version of their text. Additionally, participants chose their fictional names to report their data in the study. The second section of the questionnaire approached participants’ experience of using ChatGPT as an AWE and their views on the system’s limitations and functionalities. Examples of these questions are the open-ended questions: “Consider the commands (prompts) that you wrote to ChatGPT. Which type of prompt led to better feedback?”¹⁰ and “Consider the feedback given by ChatGPT. Is there any feedback that you liked more?”¹¹ See Appendix F for the complete questionnaire.

The third section repeats the same questions from the second section of the pre-task questionnaire, aiming to identify whether students would report any changes in their perspectives and attitudes after the experiment (for further understanding, check the definition of behavioral and attitudinal questions at the beginning of this subsection or Appendix F).

⁹ “Shared links are a new feature that allow users to generate a unique URL for a ChatGPT conversation, which can then be shared with friends, colleagues, and collaborators. Shared links offer a new way for users to share their ChatGPT conversations, replacing the old and burdensome method of sharing screenshots” (OpenAI, 2023).

¹⁰ Original question: “6. Considere os comandos (*prompts*) que você escreveu para o ChatGPT. Qual tipo de prompt levou a um melhor feedback? (Appendix F).

¹¹ Original question: “7. Considere o feedback dado pelo ChatGPT. Teve algum feedback que você gostou mais?” (Appendix F).

3.5 DATA ANALYSIS

3.5.1 Participants' Feedback Incorporation

The incorporation of feedback was operationalized using Loewen's (2004) measure "uptake". Uptake means a response to feedback received that culminates in a change of linguistic features in the text. Although Loewen's original focus is on linguistic feedback, this study adopted the same reasoning to also evaluate the incorporation of content-related feedback, given that participants requested both types of feedback from ChatGPT. To analyze feedback uptake, the researcher compared the first and second versions of the essay produced by participants, considering the feedback participants requested and received from ChatGPT.

This analysis involved examining the differences between the corresponding paragraphs from the first to the second version of participants' texts. When a modification was identified, the researcher searched the corresponding participant interaction record with ChatGPT to verify if there was a prompt made by the participant that could have led to this change. Once the prompt was identified, the prompt and the answer given by ChatGPT were compared to the outcome in the final essay to determine whether ChatGPT's output was incorporated into the final version. If there was an incorporation, the excerpts were compared to check if words or sentences were copied from the feedback into the second (and final) version of their essay. If this happened, the copied sentences/phrases/clauses were highlighted to indicate they were copied from the feedback received from ChatGPT and pasted into the final text. If ChatGPT's output excerpts were not identified in the final text, the researcher analyzed how ChatGPT's feedback influenced modifications to the final text by analyzing what was suggested by the chatbot and what had changed from the pre- to the post-feedback essay. This process allowed the researcher to determine how participants incorporated feedback in their writing tasks.

After analyzing each interaction, 3 different profiles of ChatGPT users emerged, and the researcher categorized each participant taking into consideration their interaction with the tool. A thorough definition of these groups is given in the Results and Discussion (topic 4) chapter.

3.5.2 Participants' Perceptions of ChatGPT as an AWE

The analysis of participants' reported perceptions on the topics covered in this study consisted of the comparison of the pre- and post-task questionnaires administered before and after the feedback task, respectively. This analysis aimed to elucidate whether participants reported any changes in their perceptions about the use of AI as an AWE tool after the training session and the feedback task, and also attempted to clarify why they acted the way they did when they were working on the final version of their essay with the help of the feedback received from ChatGPT.

Data analysis consisted of tabulating the questionnaires' answers into a spreadsheet. Participants' answers were grouped according to the group they were assigned to (as mentioned in section 3.5.1 and further detailed in section 4). Then, pre- and post-task answers were compared to understand how participants' reported opinions changed. These comparisons were then analyzed, taking into consideration the way they had incorporated ChatGPT feedback on their texts.

Data from sections 1 ("Personal data") and 2 ("Previous experiences with AWE tools and ChatGPT") of the pre-task questionnaire were described in full. For the section dedicated to "Beliefs regarding ChatGPT" (corresponding to section 3 in the pre-task questionnaire and section 2 in the post-task questionnaire), four questions were elicited to be discussed in the data analysis. This choice was made considering which questions could better elucidate participants' reasoning and outcomes during the feedback task. Those were:

- Three Likert-scale questions, namely:
 - "ChatGPT can be used to partially collaborate in the writing of an academic text (for example: generating topic suggestions, reviewing the text's grammar, rewriting sentences)" (Appendix G - Data sheets);
 - "It is acceptable to copy parts of the answers given by ChatGPT without referencing the source" (Appendix G - Data sheets);
 - "Understanding how ChatGPT and other Artificial Intelligence tools work is personally important to me" (Appendix G - Data sheets).
- One open-ended question: "In your opinion, how and when should the teacher talk to students about ChatGPT in their classes?" (Appendix G - Data sheets).

The procedures used for data analysis allowed the researcher to understand how the intervention — the training session and the rewriting of their essay — impacted participants' reported perceptions of ChatGPT as an AWE, elucidating their feedback incorporation process when using the tool. In the next section, the data resulting from this analysis will be discussed in the light of correlational studies on L2 acquisition, feedback, and GenAI studies.

4. RESULTS AND DISCUSSION

This section will be organized into four subsections. First, each participant from the sample is introduced together with their interaction with ChatGPT to obtain feedback and how they employed the feedback received (section 4.1). Subsequently, the section reports the analysis of the interactions of these participants with ChatGPT during the feedback task, categorizing participants' interactions in broader groups (detailed in section 4.2). After, each participant's responses to the pre- and post-task questionnaires (sections 4.3 and 4.4, respectively) are detailed. Finally, each group is examined to highlight the nuances of each participant and the contrasts that emerged within each group (section 4.5).

4.1 PARTICIPANTS' INTERACTIONS WITH CHATGPT

Participants interacted with ChatGPT in various ways, employing different strategies through their prompts. Charlie, a 31-year-old Executive Secretarial science student, debated public education workers' right to go on strike as a strategy to achieve better conditions for public education. When interacting with ChatGPT, they¹² used prompts as "please revise the following text: [...]", pasting their entire essay in the chat. ChatGPT returned with a completely corrected text, in which Charlie asked again to make it sound more academic, besides including "[...] some data and sources that reinforce the preexisting statements" (sic). Further, they also asked the chatbot to understand that the text was written by a native Brazilian Portuguese speaker using English as an L2.

Charlie's pre- and post-feedback texts differ widely in length (their second version has 173 words more than the first one). However, Charlie did not incorporate any of ChatGPT's feedback provided during the interaction. Instead, they edited their text with their own words while chatting with ChatGPT, but mostly disregarded its output. The only modification made based on ChatGPT feedback was incorporating a

¹² In this thesis, I use the singular "they" to refer to participants, as information about their gender was not covered in this study. According to the APA (2022), "Use of the singular "they" is endorsed as part of APA Style because it is inclusive of all people and helps writers avoid making assumptions about gender".

fake reference given by the chatbot, a result of an AI hallucination (Walters; Wilder, 2023).

Ash, an English Language undergraduate student, discussed the decision of Wizards of the Coast (WotC, a gaming studio company) to stop publishing Dungeons & Dragons and Magic: The Gathering materials in Portuguese and the outcomes of that decision in the Brazilian market and player base. In their interactions, Ash asked ChatGPT to “Revise the following text to be more indignant [pastes the entire text]”. ChatGPT provides a new version using expressions such as “staggering display of corporate indifference” and statements as “WotC’s actions betray their true allegiance: profit”. Later, Ash asked ChatGPT to give them “feedback to make the following text more indignant [pastes the entire text]”. This time, ChatGPT provides feedback organized in four main topics: Amplify Emotional Language, Emphasize Consequences, Direct Condemnation, and Exaggerate for Effect. The chatbot provides insights for how to improve the text within these main categories and, in the end, generates a new version of the text considering the feedback provided.

When submitting their final version, Ash copies the first text provided by ChatGPT in their interaction and pastes it. In the title, they put a note: “Magic: The Gathering despite Wizards of the Coast (v4 chatgpt, revised by me)”. Besides having used the text provided by ChatGPT, Ash highlights that they have supervised the chatbot’s work by revising it. However, the text submitted differs from the first version written by Ash. In this case, even though the main ideas presented in the text were elaborated by Ash, the second version of the essay differs in vocabulary and tone, which indicates a process of transference of authorship (Woudenberg; Ranalli; Bracker, 2024) from Ash to ChatGPT.

Bobby, a Biology undergraduate student, wrote about their experience with birdwatching and how their participation in a volunteering program led them to explore nature and birds. The participant used ChatGPT as a writing coach, starting the interaction by asking if ChatGPT thought their text was a good essay, and then pasting the entire text into the chat. ChatGPT replied with specific suggestions for each paragraph rather than correcting the entire text. As an example, ChatGPT suggested the following about their introduction: “Introduction: The introduction could be strengthened by providing a clearer hook to draw in the reader. You could start with a compelling anecdote or a thought-provoking question related to birdwatching to grab the reader’s attention”. Bobby, then, asked for clarification: “what could be an

compelling anecdote or a thought-provoking question?" (sic). ChatGPT then explained those terms, providing examples.

After ChatGPT's suggestions to their Introduction, Bobby added more information to that section of the final version of their text, explaining that their interest in a volunteering program started from the forest fires that occurred in Brazil, Argentina, and Paraguay in 2021. They repeated the same strategy for the following paragraphs, asking for suggestions on how to improve each part of the text specifically. Lastly, they asked ChatGPT, "What do think about it now?", and submitted the entire revised text again. This time, even though ChatGPT provided specific directions on how to improve, Bobby stopped modifying it and submitted it to the questionnaire. The last version of the text contained several grammatical errors (highlighted by ChatGPT in their interaction), but had no excerpts copied from ChatGPT. Bobby used the chatbot to guide their writing process, working on the individual feedback items elicited by the system.

Casey, a Librarianship student, discussed the misuses of Artificial Intelligence, its potential harms and benefits, and how to cope with such challenges. In their text, they argue that widespread use of AI should not be allowed due to the danger of it being used as a tool for oppression in Capitalism. Moreover, they address AI's potential to increase misinformation, and how inequality in underdeveloped countries may "erode opportunities for genuine human interaction". In the last paragraph, they highlight the importance of establishing boundaries and being critical of using AI.

In Casey's interaction with ChatGPT, they used a single prompt: "Make the following text more clear and cohesive. [Pastes entire text below]". ChatGPT answered with a complete overhaul of the essay, rewriting it entirely. In their final text, most of the text submitted was written by ChatGPT. Casey made minor modifications in sentences, such as: instead of "widespread and unrestricted use of innovations like AI **is questionable**", Casey preferred to write that "widespread and unrestricted use of innovations like AI **should not yet be allowed**" (Appendix G - Data sheets, emphasis added). This happened in all paragraphs, with Casey modifying parts of the ChatGPT-written output. In the final paragraph, they changed the order of the sentence written by ChatGPT, maintaining the sentence's main point. They changed it from "What's imperative is rigorous study of these tools, the development of educational programs, and the establishment of clear guidelines for their use [...]" (ChatGPT) to "The rigorous study of these tools is imperative. Educational programs must be developed, so that

people can learn to use it within clear guidelines [...]" (Casey edited, Appendix G - Data sheets).

Bennie, a German Language undergraduate, discussed how Attention Deficit and Hyperactivity Disorder can be underdiagnosed in women and the effects of it on their lives. In their essay, they argue that girls can compensate for the hardships caused by not being able to concentrate properly, getting either average or even good grades in school despite not being able to study in a practical, focused way. This would mask some of the ADHD outcomes that are usually the main signs for identifying the disorder, such as bad grades, for instance. They discuss how harmful it is for women to live with an undiagnosed disorder, which is, at the same time, overdiagnosed in men.

In their interaction with ChatGPT, Bennie prompted the following request: "Please give me feedback on how I can improve the following text: [pastes entire essay]" (Bennie, Appendix G - Data sheets). ChatGPT then provides feedback on general topics: Clarity and Conciseness, Precision, Specificity, Grammar and Punctuation, Sentence Structure, and Conclusion. They address how each of the listed attributes can be changed to improve the essay. In the end, they also provide a revised version incorporating such suggestions, without Bennie asking for it.

From the listed attributes brought by ChatGPT, Bennie employs two of the suggestions. Rather than "In the past few years" (first version of the essay), they chose to go for "In the last decade", as ChatGPT suggested to delimit a time range on their statements. Moreover, when it comes to the Specificity feedback, ChatGPT suggests "providing statistics or studies to support your claim". Bennie inserted two indirect quotations at the end of their first paragraph, mentioning that "ADHD in women is oftentimes underdiagnosed, like research has already shown (Mowlem *et al.*, 2018, Waite, 2010)" (sic). As ChatGPT did not provide these references in their interaction, I suppose that Bennie searched for articles that could back up their thesis on ADHD diagnosis in women. They did not paste the complete reference for us to confirm, but the mentioned authors are real — and not AI hallucinations (Athaluri *et al.*, 2023), as it happened in the case of Charlie — and have published papers on this field. Bennie used ChatGPT as a writing assistant, using the topics provided to make minor improvements on their essay draft.

In the following prompt, Bennie asks ChatGPT, "Which sentences could be improved for flow and readability?". ChatGPT lists seven sentences and pairs them

with a revised answer. Bennie just adopts one of them: in their first draft, she wrote, "Intelligence can also make ADHD harder to detect" (Appendix G - Data sheets). ChatGPT suggests changing "detect" to "spot". In their final text, she writes: "High intelligence can also contribute to making ADHD harder to spot".

Alex, a Law undergraduate student, discussed the misuse of amendments of judgment, a procedural instrument that delays the resolution of processes. They discuss how this instrument impacts people's lives and the need to reformulate procedural rules to prevent abuses. When interacting with ChatGPT, Alex asked ChatGPT to "Make an academic feedback about the following text: [Pastes entire text below]". ChatGPT provided feedback on specific topics, such as Clarity and Coherence, Argumentation, Grammatical Accuracy, Elaboration of Arguments, and Exploration of Potential Solutions. In the subsequent interactions, Alex asked ChatGPT to "take the recommendations of improvements that u give, and improve the text maintaining the meaning [sic]", and "Put proper academic references and 'jurisprudencia' [...]". After that, Alex asks ChatGPT to "Reformat this text respecting the ABNT norms". ChatGPT subdivided the text into sections (Introduction, Misuse and Consequences, Proposed Solutions, and Conclusion). Finally, Alex prompts ChatGPT to "exclude the topical format, make it a continuous text (still preserving ABNT norms)". ChatGPT generates the text without the topics, and Alex copies the last version generated by ChatGPT and submits it.

Blake, who is also a Law student, discussed the internet's role in political radicalization and the rise of populist politicians. During their interaction, they prompted ChatGPT by asking for "feedback on the following text," providing his entire essay. ChatGPT offered structured feedback and suggested expanding on the global perspective and making the call to action more specific. ChatGPT also recommended a more balanced tone. Blake selectively applied ChatGPT's feedback, adapting certain parts of his text without incorporating all modifications. For instance, he added specific examples of global figures and clearer actions for his call to action, demonstrating a nuanced approach to the AI's suggestions.

4.2 FEEDBACK ANALYSIS

After comparing pre- and post-task essays with participants' interactions with ChatGPT, patterns were identified regarding the incorporation of ChatGPT's feedback. Participants were divided into three groups based on their attitudes regarding ChatGPT feedback: A - integral incorporation or plagiarism; B - writing assistance; and C - partial incorporation (Appendix G for datasheets).

Table 1 – Groups divided by number of participants and the type of feedback incorporation

Group	Number of participants	Type of incorporation	Participants ¹³
A	2	Integral incorporation	Ash, Alex
B	3	Writing assistance	Bobby, Blake, Bennie
C	2	Partial incorporation	Casey, Charlie

Source: the author, 2025.

In group A, students incorporated the new text provided by ChatGPT without modifications. They used prompts such as “Revise the following text [...]” and “Make academic feedback about the following text: [...]”. They transferred the entire writing process to AI rather than incorporating their modifications within the text (Table 2). This type of incorporation was incompatible with the orientations in the training session, which occurred minutes before the feedback task. Figure 4 presents one of the slides from the slideshow used in the session presenting acceptable and unacceptable attitudes toward incorporating LLMs in writing, according to Laquintano, Schnitzler, and Vee (2024). Even though participants were advised that they should not “transfer the main part of the writing process and research to AI,” participants from group A did so.

¹³To protect participant anonymity while avoiding impersonal labels (e.g., Participant 1, Participant 2), gender-neutral pseudonyms were assigned. Each pseudonym begins with the letter corresponding to the participant's group (A, B, or C) to support clearer interpretation of the results and easier identification of group affiliation.

The table below compares the corresponding excerpt in the first version of the essay, in the modified text provided by ChatGPT's output, and in the second version of the essay submitted by participants. The ChatGPT feedback column and the final version column are identical, as group A participants integrally incorporated ChatGPT's output.

Table 2 – Excerpts demonstrating integral incorporation by Group A

Participant	Excerpt from the first version	Excerpt from ChatGPT's feedback	Excerpt from the second version
Ash	The Brazilian player base obviously loves WotC's games, but at every turn WotC makes it clear that their only love is money. So MtG will go on in Brazil, not because of WotC, but in spite of it.	Despite the unwavering devotion of the Brazilian player base, WotC's actions betray their true allegiance: profit. MtG may persist in Brazil, but not due to WotC's benevolence, rather in defiance of their corporate greed and disdain.	Despite the unwavering devotion of the Brazilian player base, WotC's actions betray their true allegiance: profit. MtG will persist in Brazil, but not due to WotC's benevolence, rather in defiance of their corporate greed and disdain.
Alex	The Brazilian law have many Processualist instruments that assure the legal process and proporcionate the contradictory. However, one of this instruments, especilally the "Embargos de Declaração", has mostly used to procrastinate the time of the law and making the Brazilain justice slower and tedious.	The Brazilian legal system boasts numerous procedural instruments designed to ensure due process and uphold the principles of fairness and equity. However, one such instrument, the "Embargos de Declaração," has regrettably become a tool often wielded to delay proceedings, leading to a cumbersome and protracted justice system.	The Brazilian legal system boasts numerous procedural instruments designed to ensure due process and uphold the principles of fairness and equity. However, one such instrument, the "Embargos de Declaração," has regrettably become a tool often wielded to delay proceedings, leading to a cumbersome and protracted justice system.

Source: the author, 2025.

Group B used ChatGPT as a writing assistant. They used the chatbot to ask for tips, suggestions, and ideas on text improvement. They did not incorporate ChatGPT's suggestions integrally, as they interacted with ChatGPT to provide assistance and not to reformulate their text. Group B interacted with the system to create iterative feedback based on suggestions given before. They used prompts like "Do you think this text is a good essay [pastes the entire essay]", "Please give me feedback on how can I improve the following text: [...]"; and "can you give me some feedback on the following text?" (in this last example, participant pasted their essay after ChatGPT's response, to their initial prompt).

In all interactions from group B, ChatGPT provided feedback organized into topics. In Bobby's interaction, ChatGPT provided feedback for each paragraph. At the same time, Blake and Bennie received feedback on broader topics, as "Clarity and Conciseness", "Grammar and Punctuation", and "Clarity and Structure", for instance. The table below presents an excerpt from the first version, the feedback provided by ChatGPT to that excerpt, and the second version of the excerpt submitted by participants. Text highlighted in green indicates excerpts that were maintained from the first version of the essay, and text highlighted in yellow represents modifications based on ChatGPT output.

Table 3 – Excerpts demonstrating Group B's interactions

Participant	Excerpt from the first version	Excerpt from ChatGPT's feedback	Excerpt from the second version
Bobby	All of it begins in the middle of the pandemic in 2021. After the quarantine was a little more flexible, I was lost in what to do with my life. I came across an ong from my country that was	This text has the potential to be a good essay, but it could benefit from some improvements to enhance clarity, coherence, and depth. Here are some suggestions: Introduction: The introduction could be strengthened by providing a clearer hook	All of it begins in the middle of the pandemic in 2021. After the quarantine was a little more flexible, I was lost in what to do with my life, like probably teh mejority of the population in the world, mourning for something that would give my life some porpose. Something worth metioning is that at the same time there was a lot of

	offering volunteer work in a protected forest area, I signed for it and waited the results.	to draw in the reader. You could start with a compelling anecdote or a thought-provoking question related to birdwatching to grab the reader's attention. [...]	forest fires around Brazil, Argentina, and my country Paraguay that just was depressing. Some day, swipping through Instagram I came across an ong from my country that was offering volunteer work in a protected forest area, I thought to myself "Would these makke me feel like Im doing something to help at least?" and at the time I think it was doing the bare minimun but even doing that helps and also the motivation to meet new people and do new activities , I signed for it and waited the results. [...]
Blake	As it stands, this phenomena has been bargely replicated in other countries like Brazil or Argentina, and the damages it has caused are already tangible, it is the time for us to rally against this self-destructing culture for the well-being of man-kind, we either do it now or forever remain trapped in then bigotry.	Here are some feedback points on your text: [...] Global Perspective: You briefly mention that similar phenomena have been observed in other countries like Brazil and Argentina, but you could expand on this point. Providing more context or examples from other countries would strengthen your argument and demonstrate the global nature of the issue. Call to Action: The text ends with a call to action, urging readers to rally against the perceived negative impacts of right-wing ideologies. However, it could be more specific about what actions readers can take to address the issue, whether it's through political engagement, community	As it stands, this phenomena has been bargely replicated in other countries like Brazil or Argentina, figures like Javier Millei and Jair Bolsonaro taking Trump's place in the radicalization, and the damages it has caused are already tangible, it is the time for us to rally against this self-destructing culture for the well-being of man-kind, we have to take our voice to outlets that will make them heard and try to reduce the reach of the radicalizing outlets. We either do it now or forever remain trapped in the bigotry.

		organizing, or other means. [...]	
Bennie	In the past few years, the discussions around the topic of Attention Deficit and Hyperactivity Disorder seem to have doubled or even tripled. More specifically, people have begun to notice that ADHD in women appears to be oftentimes underdiagnosed.	Your text is well-written and effectively communicates the issue of underdiagnosis of ADHD in women. Here are some suggestions for improvement: Clarity and Conciseness: Consider tightening up the language for clarity and conciseness. For example, you could streamline the first sentence to something like: "In recent years, discussions about Attention Deficit and Hyperactivity Disorder (ADHD) have markedly increased." Precision: Specify the time frame when discussing the increase in discussions about ADHD. Instead of "the past few years," you could mention a specific range, such as "in the last decade" or "since the early 2000s." Specificity: When discussing the underdiagnosis of ADHD in women, provide statistics or studies to support your claim. Adding specific data can strengthen your argument and provide credibility to your assertions. [...]	In the last decade, the discussions around the topic of Attention Deficit and Hyperactivity Disorder (ADHD) seem to have doubled or even tripled. More specifically, the general public has begun to notice that ADHD in women is oftentimes underdiagnosed, like research has already shown (Mowlem et al., 2018, Waite, 2010). [...]

Source: the author, 2025.

In Group C, Casey and Charlie prompted ChatGPT to “Make the following text more clear and cohesive [pastes entire essay]” and to “[...] revise and correct the following text [essay]”. In both cases, ChatGPT's output was a thoroughly revised text with changes. Both participants asked ChatGPT to review the text entirely, as Group A did, but incorporated some of its suggestions to assist their writing, as Group B did. What distinguishes group C is that they copied entire excerpts from ChatGPT, rather than the entire text generated by ChatGPT.

Table 4 shows excerpts from group C participants to illustrate how this process occurred. Text highlighted in green represents excerpts that were kept from the original text without modifications. Text highlighted in yellow represents excerpts that were copied from ChatGPT. Text in orange represents excerpts from ChatGPT that were adapted into the new version.

Table 4 – Excerpts demonstrating Group C partial incorporation

Participant	Excerpt from the first version	Excerpt from ChatGPT's feedback	Excerpt from the second version
Casey	Since undevelopped countries cannot hope to accompany the speed of change in technology, it cannot be expected that they adopt to work with such tools easily. This alienation and disparity created and taken advantage of by capitalism eliminates the time and space for human interation, originality of thought and expression, and autonomy. Therefore, it is not yet safe for tools like AI to be used without limitations, without debate or	Given the inability of underdeveloped countries to keep pace with rapid technological advancements, expecting them to seamlessly integrate such tools into their systems is unrealistic. This alienation and disparity, perpetuated and exploited by capitalist forces, erode opportunities for genuine human interaction, original thought, and autonomy. Thus, it's premature to embrace tools like AI without constraints, discussions, or critical analysis, as they can inflict harm on individuals and communities. What's imperative is rigorous study of	Since undevelopped countries cannot hope to accompany the speed of change in technology, to expect them to seamlessly integrate such tools into their systems is unrealistic. This alienation and disparity, created and exploited by capitalist forces, erode opportunities for genuine human interaction, original thought, and autonomy. Thus, it's premature to embrace tools like AI without stablishing boundaries, discussions, or critical analysis of it's uses and functions, as they can be used to inflict harm on individuals and

	critical thought because people can suffer from it. What needs to be done is study of said tools, the creationg of teaching programs and usage with objective and rules, so it won't be used without caution.	these tools, the development of educational programs, and the establishment of clear guidelines for their use, ensuring they are employed with caution and accountability.	communities. The rigorous study of these tools is imperative. Educational programs must be developed, so that people can learn to use it within clear guidelines, ensuring they are employed with caution and responsibility.
Charlie	As a basic Human Right, Education should not only be public and available, it must have quality and accessibility. If professors and administrative workers are not being rightfully compensated, the result can be unmotived and sick professionals, and the decrease in quality is almost unavoidable.	As a fundamental human right, education must not only be publicly accessible but also uphold standards of quality and inclusivity. Research indicates that inadequate compensation and working conditions among academic and administrative staff not only lead to professional dissatisfaction but also contribute to declines in educational quality (Jones & Brown, 2018). Consequently, the failure to provide fair compensation risks demotivating and impairing the well-being of professionals, inevitably compromising educational standards.	As a basic Human Right, Education not only should remain public and available, it also needs to have quality and accessibility assured. Evidence suggests that inadequate compensation and working conditions among academic and administrative staff not only generates less satisfied professionals, but also generates declines in educational quality (Jones & Brown, 2018). Therefore, if professors and administrative workers are not being rightfully compensated, the result can be unmotived and ill professionals, and the decrease in quality is almost unavoidable.

Source: the author, 2025.

Even though they are in the same group, Casey and Charlie had different behaviors while partially incorporating ChatGPT's feedback. Casey used most of the ChatGPT-generated version of their text and adapted some excerpts. On the other hand, Charlie did not incorporate most of the feedback from ChatGPT. Mostly, the modifications made in their text were all made by the participant, and not suggested, written, or edited by ChatGPT. However, Charlie incorporated an AI-generated fake bibliographic reference provided by ChatGPT, that is, an AI hallucination (Walters; Wilder, 2023).

In the following subtopics, each participant's pre-task and post-task questionnaires will be analyzed, and their attitude towards ChatGPT as a source of feedback will be compared.

4.3 PRE-TASK QUESTIONNAIRE

This section describes the answers to the pre-task questionnaire, administered before the training session and the feedback task.

4.3.1 Group A - Integral incorporation

Alex is a Law student and reported that they had never studied English formally before this course. They had experience with AWE tools, such as Microsoft Word and Google Docs' automatic spell checkers and online dictionaries, and sometimes used ChatGPT to test it (less than once per month). They had never used ChatGPT for a writing assignment in a Foreign Language but had already used the chatbot for other academic assignments with the objective of formulating "a text quickly and easily, using the information they had inserted"¹⁴. They defined ChatGPT as "A text generator that uses information from the internet as a database."

On their beliefs regarding ChatGPT, Alex partially disagreed that ChatGPT can be used to partially collaborate with writing an academic assignment and strongly

¹⁴ Translated from the original, in Portuguese: "O objetivo era formar um corpo textual de maneira rápida e prática, com as informações que eu inseri. Sim." (Data sheets, Appendix G).

disagreed that it is acceptable to copy parts of ChatGPT's outcomes without referencing it. As they oppose plagiarism when using ChatGPT, Alex strongly agreed that understanding how ChatGPT and other AI tools work is important for them. When questioned on how and when professors should talk about ChatGPT in their classes, Alex wrote that at the beginning of academic courses, professors should keep an open dialogue about when and how AI tools could be used.

Ash is an English Language undergraduate student. They are 28 years old and have studied English for 22 years. They reported having had experience with some AWE tools, such as Microsoft Word and Google Docs' automatic spell checkers and online dictionaries. They had never used ChatGPT before and described it as an online tool that uses AI to generate text. They reported to agree that ChatGPT can be used to partially collaborate on writing an academic text and strongly disagreed that it is acceptable to copy parts of ChatGPT's outcomes without referencing it. They disagree that understanding how ChatGPT and other AI tools work is personally important for them — in their opinion, at the beginning of the course, it is important to highlight the relevance of academic integrity and how ChatGPT can be misused.

4.3.2 Group B - Writing assistance

Blake is a Law student who reported to have studied English for 14 years out of their 19 years. They claimed to have little experience with AWE tools, only having used Microsoft Word's automatic spell check. According to the participant, they had used ChatGPT at least once a month in the months preceding the moment they answered the questionnaire. They declared to have never used ChatGPT to perform a Foreign Language task, though they had already employed it in writing academic assignments. Even though they did not share what their objectives were with ChatGPT on the academic assignment, they described ChatGPT as an AI tool that draws from internet-based information and performs better with older topics, due to reinforcement from prior learning.

When it comes to their beliefs on whether ChatGPT can be used to partially collaborate on writing an academic text, Blake agreed with it. On the other hand, Blake reported strongly disagreeing with copying parts of the output given by ChatGPT without referencing it. They also declared partially disagreeing that understanding how

ChatGPT and other AI tools work is personally important for them. Blake argued that teachers should address the use of ChatGPT as a possibility for learning with students, but alert them about its shortcomings and potential harms when overused.

Bobby is a 23-year-old Biology undergraduate student who has studied English for 13 years. They have extensive experience with AWE, having used tools such as Microsoft Word and Google Docs' automatic spell checkers, Grammarly, and online dictionaries. They reported using ChatGPT every day, including for all Foreign Language tasks mentioned in the questionnaire: translating a text, writing an email, revising a text, helping out to study for a test, explaining a concept, and summarizing a text. They have also used ChatGPT to prepare slideshow presentations, to have a guide for summarizing their text, and to help them better understand concepts. For them, the prompts used accomplished their objectives.

Bobby defined ChatGPT as a tool that derives its answers from its database. It is stimulated by and keeps learning through its interaction with users. Regarding their beliefs, Bobby reported to agree that ChatGPT can be used to partially collaborate on writing an academic text. They disagreed that it is acceptable to copy parts of the excerpts given by ChatGPT without referencing them. Moreover, Bobby partially agreed that understanding how ChatGPT and other AI work is important to them. Lastly, they argued that professors should approach how to use ChatGPT to clarify themes and concepts and not prohibit its use.

Besides being a German Language undergraduate student, Bennie has studied English for 13 years. They are 24 years old and have experience with AWE tools, such as Google Docs' automatic spell check, Grammarly, and online dictionaries. In the preceding months of the study, they used ChatGPT only to test it (less than once per month), and have used it for some Foreign Language-related tasks, such as translation, explaining a concept, and summarizing a text, though they mentioned having never used ChatGPT for an academic task.

Bennie defined ChatGPT as a tool capable of summarizing texts as well as translating, producing, and revising documents. They reported agreeing that it can be partially used for a writing assignment and strongly disagreed that it is acceptable to copy ChatGPT's output without referencing it. They also partially agreed that understanding ChatGPT and AI is important for them. For Bennie, professors should approach ChatGPT in their classes, aiming to show students its ability to quote and create wrong or nonexistent references and to manipulate data to achieve the prompt's

objective. They also mentioned that professors should teach students how to recognize texts created by ChatGPT and how to use the tool for good purposes.

4.3.3 Group C - Partial incorporation

Casey is a 23-year-old Librarianship undergraduate student who has studied English for 10 years. They have previous experience with Microsoft Word and Google Docs' automatic spell checkers, but had never used ChatGPT. They defined it as an AI tool created to facilitate specific text, image, and other digital formats writing and editing tasks. Casey reported partially disagreeing that ChatGPT can be used to collaborate on writing academic texts and strongly disagree that it is acceptable to copy excerpts from ChatGPT's output without referencing it.

Moreover, Casey partially disagreed that understanding how ChatGPT and other AI tools work is essential for them. When questioned about how and when professors should approach ChatGPT use in their classes, they argued that Academia should discuss ChatGPT's implications. However, they mentioned that it should not be used in basic education due to the "lack of knowledge and critical thought of teachers and students"¹⁵.

Charlie is 31 years old and has studied English for over 10 years. They are an Executive Secretarial science undergraduate student and use Microsoft Word and Google Docs' automatic spell checkers and online dictionaries when writing. They have used ChatGPT less than once per month and affirm they have used it to perform Foreign Language-related tasks — to explain a concept and brainstorm, specifically. They also affirmed they have used ChatGPT on academic tasks, such as preparing slideshow presentations and finding topics related to a particular matter.

Regarding their experience using ChatGPT for academic tasks, Charlie affirmed that their objectives were "to have an initial outline as a starting point to correct and improve [the text] and to have insights for topics and arguments. I was able to

¹⁵ Adapted and translated from the original, in Portuguese: "acho correto que o ambiente universitário debata seus usos, mas não que seja usado em escolas de ensino básico, por falta de conhecimento e pensamento crítico de professores e alunos" (sic). (Data sheets, Appendix G).

reach those objectives partially because the tool does not offer references”¹⁶. Charlie defined ChatGPT as a tool that has an extensive database and is capable of answering and formulating complex texts automatically based on the user's input.

In regard to their beliefs on ChatGPT, Charlie reported to strongly agree that ChatGPT can be partially used when writing an academic text. They partially agreed that it is acceptable to copy parts of ChatGPT's output without referencing it. Charlie also strongly agreed that understanding how ChatGPT and other AI tools work is important for them. When asked about how and when professors should approach ChatGPT in their courses, Charlie affirmed that at the beginning of the course, professors should warn students about plagiarism, the lack of references, originality, and the use of entire excerpts copied from ChatGPT's output.

4.4 POST-TASK QUESTIONNAIRE

4.4.1 Group A - Integral incorporation

Alex reported that they liked the grammar and formatting feedback given by ChatGPT and that there was no piece of feedback they disliked. They reported that they would use ChatGPT again in other academic activities, but would not let ChatGPT write for them again. Alex indicated that they would only consider punctuation corrections given by ChatGPT, but would correct it by themselves. In the post-task questionnaire, Alex now reported to agree that ChatGPT can be used to partially collaborate on writing an academic text (they partially disagreed before the feedback task). They still reported to strongly disagree that it is acceptable to copy parts of ChatGPT's output without referencing it, and they reported to still strongly agreeing that understanding how ChatGPT and other AI tools work is personally important for them.

Finally, in the post-task questionnaire, Alex presented a more robust opinion on how and when professors should talk about ChatGPT in their classes: according to Alex, when a text is overly generic or when formatting is so “perfect” that it looks

¹⁶ Translated from the original, in Portuguese: “[...] ter um esboço inicial, como ponto de partida, para depois fazer correções e aperfeiçoar, e ter ideias de tópicos e argumentos. Consegui alcançar esses objetivos parcialmente porque a ferramenta não oferece referências” (Data sheets, Appendix G).

robotized. In their opinion, these are some points that professors should indicate to their students to signal the lack of personality and authorship in a piece of writing. This suggests that, for Alex, professors should address ChatGPT and AI chatbots only after students have used the tool, as no strategy for its use has been indicated.

Ash reported not having specific feedback given by ChatGPT that they liked more. They mentioned that they would avoid using ChatGPT again in other academic tasks, because they believe they would not use it ethically all the time. Ash maintained their beliefs regarding ChatGPT after the training session and the feedback task: they reported to still agree that ChatGPT can be used to collaborate on an academic task partially, strongly disagree that it is acceptable to copy parts of ChatGPT's output without referencing it, and disagree that understanding how ChatGPT works is important for them.

After the feedback task, Ash demonstrated a more complex view of how and when teachers should approach AI and ChatGPT. According to them, professors should present ethical considerations about using AI to produce academic texts, as this would prevent further “problems [sic]”. After this, professors could teach how to use AI tools to improve texts produced in and for classes.

4.4.2 Group B - Writing assistance

Blake reported that, in general, the feedback given by ChatGPT was helpful. They do not report any piece of feedback they did not like. Blake reported that they would use ChatGPT again on other tasks and would not change anything from the approach they had on the feedback task. Previously, they reported to agree that ChatGPT could be used to collaborate on academic tasks, whereas now they strongly agree. Also, now they disagreed that it is acceptable to copy parts of ChatGPT's output without referencing it — beforehand, they strongly disagreed with this affirmation. Still, they now partially agreed that understanding how ChatGPT works is personally important for them — a relevant change from the pre-task questionnaire, where they partially disagreed with it.

When asked about how and when professors should approach ChatGPT, Blake sounds more factual than in the pre-task questionnaire: they argued that professors should approach it honestly and directly because they cannot prevent

students from using it. Thus, according to Blake, the most efficient way is to appeal to the goodwill of students when it comes to the use of AI tools.

Bobby reported having liked ChatGPT's output when they asked it to give examples of possible improvements and/or changes. Bobby disliked that ChatGPT rewrote their text without being asked to do so. They mentioned they would use ChatGPT again in other academic tasks but would avoid using ChatGPT's vocabulary. Regarding their beliefs on ChatGPT, Bobby reported to partially agree that ChatGPT can be partially used to collaborate on an academic assignment — an affirmation they agreed with before.

Moreover, Bobby now reported to partially disagree that it is acceptable to copy parts of ChatGPT's output without referencing — a statement they would disagree with in the pre-task questionnaire. Bobby demonstrated being more interested in ChatGPT after the task: they reported to agree that understanding how ChatGPT works is important for them, something they partially agreed with before. Now, when asked about when and how professors should approach ChatGPT, Bobby presented a more elaborate perspective: they say professors should instruct the ways in which ChatGPT should be approached and not prohibit or stigmatize its use.

Bennie did not like it when ChatGPT suggested that their text be shorter and that sentences be simplified. They would not use ChatGPT again in other academic tasks and would not do anything different from what they did. Regarding their beliefs, Bennie reported now to partially agree that ChatGPT can be partially used to collaborate on academic texts, something they agreed with before. In the pre-task questionnaire, Bennie reported to strongly disagree that it is acceptable to copy parts of ChatGPT's output without referencing it — something they reported partially disagreeing with now. Finally, Bennie demonstrated to be less interested in understanding how AI and ChatGPT work after the training session. Before, they reported to partially agree that understanding how AI tools work is important for them; after the feedback task, they partially disagreed with it.

When asked about how and when professors should approach ChatGPT, Bennie argued that professors should talk about ChatGPT, require writing assignments only written by students without any help of AI, and talk about what is considered acceptable or not when using AI on assignments. Moreover, professors should mention how to use it cautiously.

4.4.3 Group C - Partial incorporation

“It turned the text more reader-friendly than my original [version]”¹⁷, says Casey when asked about what feedback given by ChatGPT they liked the most. On the other hand, Casey argues that they did not like it when ChatGPT changed essential parts of their text. They argue ChatGPT changed their positioning from “AIs still should not be used with such freedom” to “the use of AIs is still questionable”. Casey also perceived an exaggerated replacement of words by synonyms, which led them to maintain some original excerpts from their text in the final version of the essay.

Casey reports they will not use ChatGPT for any academic, scientific, or personal task. They mention they would prefer if they had had more guided, in-class experiences when using ChatGPT’s feedback in their text, rather than using it “freely”. They also complain about the conditions under which the feedback task was conducted.

As I relied on my memory, the concepts that I usually have in mind to write a quality text flew away in this activity, maybe due to my lack of habit in using ChatGPT to evaluate, adapt, and edit my text, maybe due to the distinct nature of what I consider the act of writing itself¹⁸.

Casey disagrees that ChatGPT could be partially used in an academic task, an affirmation that they partially disagreed with before the feedback task. They still strongly disagree that it is acceptable to copy excerpts from ChatGPT output without referencing it. Casey is more interested in how AI and ChatGPT work after the feedback task and the training session. Before, they partially disagreed that understanding AI is important for them, something they now partially agree with.

Once again, Casey argues that basic education teachers are not ready to comprehend, talk, or accept the use of ChatGPT in school assignments. On the other

¹⁷ Adapted and translated from the original, in Portuguese: “Em termos de organização do texto, ele o fez bem, tornou o texto mais reader-friendly que o meu original” (Casey in the post-task questionnaire, Data sheets, Appendix G).

¹⁸ Adapted and translated from the original, in Portuguese: “Como dependi da memória, os conceitos que normalmente eu tenho em mente para escrever um texto de qualidade me fugiram um pouco nessa atividade, talvez por falta de costume de usar o Chat GT para avaliar, adaptar e editar meu texto, talvez pela natureza distinta do que considero a atividade de escrita em si” (Casey in the post-task questionnaire, Data sheets, Appendix G).

hand, university professors should debate it as much as possible. For them, accepting ChatGPT or not would depend on the willingness of each professor.

Charlie did not like any feedback given by ChatGPT. According to them,

I didn't like any [piece of feedback], when it improved something, it went back and worsened other things I had already asked to change before. For example, I did not asked at any moment to change terms and expressions, and even asking explicitly not to change [the terms and expressions], it [ChatGPT] would keep doing so. And when I asked to use less pompous language, it would write "like dummy."¹⁹

After this negative experience, Charlie was unsure if they would use ChatGPT again in other academic tasks. They say it is not even useful for them to correct text, once it only confuses them more and modifies things it should not change.

Related to this negative report, Charlie decreased their level of belief that ChatGPT could be used partially in writing assignments. Before, they strongly agreed with this affirmation, whereas now they only partially agree with it. Moreover, they still partially agree that it is acceptable to copy excerpts of ChatGPT's output without referencing it, and strongly agree that it is important for them to understand how ChatGPT works. Finally, when answering about how and when professors should approach AI in their classes, Charlie maintains her argument that professors should always highlight issues such as plagiarism, originality, and critical thought at the beginning of their courses, similar to what they argued in the pre-task questionnaire.

4.5 GROUPS ANALYSIS

Following the analysis of participants' interactions with ChatGPT and their responses to pre- and post-task questionnaires, this section provides a comparative analysis of the three groups identified in this study: Group A (Integral incorporation or plagiarism), Group B (Writing assistance), and Group C (Partial incorporation). Based

¹⁹ Adapted and translated from the original, in Portuguese: "nao gostei de nenhum, quando melhorava uma coisa, voltava a piorar outra que eu ja tinha pedido pra mudar antes. por exemplo, em nenhum momento pedi pra mudar termos e expressoes, e mesmo pedindo explicitamente pra não fazer, ele continuava mudando. E quando pedir ora usar linguagem menos rebuscada, ele escrevia like dummy" (Charlie in the post-task questionnaire, Data sheets, Appendix G).

on the data presented in the preceding subsections (4.1 to 4.4), this section examines the differences among participants within each group and highlights the key differences observed across the groups.

4.5.1 Group A analysis

This section presents an analysis of Group A, which comprised two participants (N=2) and was characterized by the integral incorporation of text generated by ChatGPT into their final essays. A notable discrepancy was observed between the group's stated beliefs and their actual behavior on the feedback task. While integrally incorporating ChatGPT's text into their assignments, both participants argued, in pre- and post-task questionnaires, that they strongly disagree that it is acceptable to copy ChatGPT's outcome without referencing it. Participants argued that professors should keep an open dialogue on how and when to use ChatGPT on assignments, and Ash even highlighted that academic integrity should be reinforced, mainly because ChatGPT can be misused.

The inconsistency between their expressed belief against plagiarism and their actions on the feedback task is relevant, given that group A was the only group in which both participants strongly disagreed that it is acceptable to plagiarize ChatGPT in both questionnaires, yet it was the group categorized as integral incorporation. Several factors may have contributed to this phenomenon, including the discrepancy between their beliefs and behaviors (Hu *et al.*, 2016), the lack of prior experience with AI and LLMs, and personal factors, such as their objectives when accomplishing the feedback task.

Discrepancies between beliefs and behavior have been reported in the literature, as self-discrepancy theory proposes that people behave aiming to reduce the differences between who they are and who they want to be, both in physical and virtual environments (Hu *et al.*, 2016). This type of behavior can also be influenced by social desirability bias, that is, "Due to self-presentation concerns, survey respondents underreport socially undesirable activities and overreport socially desirable ones" (Krumpal, 2011, p.2025). The strong disagreement with plagiarism, opposed to the act of copying and pasting texts entirely generated by ChatGPT, can, in part, reflect an

attempt from Ash and Alex to present a socially acceptable point of view in the questionnaires.

Moreover, other possible reasons for the inconsistency of their behavior and beliefs are their low level of AI literacy, their lack of experience with ChatGPT, and the insufficiency of the training session in developing strategies to cope with GenAI. Alex reported having used ChatGPT less than once per month before the study, and had never used it for a writing assignment in a Foreign Language, even though they used it for other academic writing assignments. Ash reported never having used ChatGPT in any context. Their limited prior experience compared to other groups might explain the integral incorporation of ChatGPT's output, as less experienced and critical students may over-rely "on AI tools to meet language standards" (Shen; Chen, 2025, p.21).

Finally, the motivation for such an attitude may have an interplay with other aspects of L2 development and writing that were not covered in this study, such as participants' motivation, emotions, and other individual differences in language learning (VanPatten; Williams, 2015). As demonstrated in one fragment of Alex's pre-task questionnaire answers, an apparent previous motivation when using ChatGPT is reported when they state that, in their prior experience with ChatGPT, their objective was to "[...] formulate a text quickly and easily from the information I provided." (Alex, Pre-task questionnaire, Appendix G). Their objectives — generating text in a quick and easy way — are associated with personal variables, as reported by Galindo-Dominguez *et al.* (2025, p.9): "The factors that truly influence the increase or decrease in plagiarism levels are other personal variables associated with students, such as cheating culture or amotivation." Therefore, further studies might consider measuring other aspects of (language) learning to intertwine human-AI interaction through chatbots and feedback incorporation.

In sum, group A highlights a contradictory and complex context of participant behavior when interacting with ChatGPT. Even though they have strongly opposed plagiarism through AI use, both participants engaged in a behavior that contradicted their stated beliefs. This gap between their intentions and actual attitudes may be explained by social desirability bias, in which respondents try to report the ideal answer rather than their true beliefs, especially when it comes to sensitive and ethical topics (Chung; Monroe, 2003). Moreover, the inconsistency of group A may occur due to a combination of personal factors (such as individual differences in language

development and AI literacy skills), but also from insufficient scaffolding and training, which could have attenuated individual aspects. The findings suggest that integrating GenAI tools (such as ChatGPT) in L2 educational settings requires more than technical instruction, as it demands from pedagogical strategies that address students' beliefs and their varying levels of language domain and critical thinking.

4.5.2 Group B analysis

This section examines Group B, comprising three participants (N=3), whose interactions with ChatGPT were characterized as writing assistance. An important distinction from this group compared to Group A is that they do not incorporate ChatGPT's output integrally into their writing assignment. Instead, they use the chatbot to request improvement on specific domains, interacting iteratively to refine their essays.

Despite having performed in similar ways on the feedback task, participants from group B have different backgrounds when it comes to their previous experience with ChatGPT. Blake had never used ChatGPT for Foreign Language tasks. In contrast, Bobby had used it for all foreign language-related possibilities mentioned in the questionnaire (translating a text, writing an e-mail, revising a text, helping to study for a test, explaining a concept, and summarizing a text). Bennie has used the chatbot for translating texts, explaining a concept, and summarizing a text. It is interesting to highlight that participants Blake and Bobby are the participants who had used ChatGPT the most in the months before answering the questionnaire: Blake used it at least once a month, and Bobby used it every day. Bennie used ChatGPT less than once per month.

As stated in subsection 4.2, Blake, Bobby, and Bennie did not incorporate ChatGPT's suggestions integrally but used some of the insights to modify their texts. This type of behavior has been previously described in the literature, as in Shen and Chen's (2025, p.20) study, where "participants selectively applied ChatGPT-generated content with justification to refine or supplement their writing, based on self-regulatory judgment according to their writing requirements and personal needs". For the authors, this behavior contradicts a common sense that the use of ChatGPT in academic

assignments can simplify or lead to more shallow cognitive processes, as trained students might further develop their critical thinking skills when interacting with GenAI.

Group B's critical behavior towards ChatGPT supported them in leveraging the chatbot output during the feedback task. This aspect, together with their higher exposure to ChatGPT before the study when compared to other participants, leads to the hypothesis that their previous experience with ChatGPT may have positively impacted their reasoning during the feedback task. In this sense, their previous knowledge may have prepared them to better cope with ChatGPT limitations and the ethical issues implied in copying excerpts from a GenAI chatbot, as the increased use of ChatGPT can positively influence the ethical use of AI by participants (Kajiwara; Kawabata, 2024).

The definitions of ChatGPT provided by group B in the pre-task questionnaire also corroborate the hypothesis of group B's apparent higher level of AI literacy (UNESCO, 2023a). Blake and Bobby used words such as "experience" and "stimulus", referring to Machine Learning concepts. At the same time, Bennie provided a simpler definition, defining ChatGPT by its capabilities of summarizing texts, translating, producing, and reviewing documents etc. Group B's apparent higher level of AI literacy led to a performance closer to what was expected in the feedback task (Appendix D and Figure 5).

The analysis of group B's feedback task and their pre- and post-task questionnaires elicited the use of ChatGPT by this group as a writing assistant, suggesting a more critical and analytical use of the chatbot when compared with group A. Group B's writing assistance use of ChatGPT demonstrates a more critical engagement with ChatGPT, highlighted by the selective use of its output rather than full reliance on content generated by AI. The group's apparent higher level of AI literacy, demonstrated by their greater previous experience with ChatGPT and the more detailed conceptualization of the feature on the pre-task questionnaire, may have positively influenced ethical and pedagogical use of the tool, both in prompt creation and feedback processes. Nonetheless, Bennie's opposition to AI instruction in basic education reveals misconceptions about literacy development. Such views highlight the need for a broader view of AI literacy, which integrates critical thinking and ethics (as discussed in subtopic 4.5.1 - Group A analysis).

4.5.3 Group C analysis

This section analyses Group C's (N = 2) performance throughout this study. Participants in this group partially incorporated excerpts written by ChatGPT into their texts. During their performance on the feedback task, they interacted with ChatGPT by requesting feedback on the entire text, similarly to the strategy employed by group A. However, unlike group A's integral incorporation, Casey and Charlie merged specific excerpts from ChatGPT's output into their original text. Casey, for instance, copied significant parts from ChatGPT's output, making minor modifications to it (Table 5). Charlie, conversely, did not use ChatGPT's feedback in most of the text. However, after modifying most of the text according to their own review, Charlie copied two sentences from ChatGPT's output, one of them being a fake reference²⁰ (Table 5).

Casey and Charlie have similar backgrounds when it comes to their previous experience with AWE tools and ChatGPT. In the pre-task questionnaire, Charlie stated that, before the study, they had used Microsoft Word and Google Docs' automatic spell checkers and online dictionaries, and Casey only used Microsoft Word and Google Docs' automatic spell checkers. Charlie reported having used ChatGPT less than once a month before the study, including for L2-related tasks, as summarizing texts and brainstorming. On the other hand, Casey reported they had never used it.

In spite of having low familiarity with ChatGPT and being categorized in the same group, Casey and Charlie have demonstrated opposite perceptions of ChatGPT during the course of the study. Starting from their perceptions on ChatGPT and academic writing on the pre-task questionnaires, Casey and Charlie partially disagreed and strongly agreed, respectively, that ChatGPT can be partially used on writing assignments. Moreover, Casey strongly disagreed and Charlie partially agreed that it is acceptable to copy parts of the answers from ChatGPT without referencing it, points of view maintained in the post-task questionnaires.

Even though Casey affirmed they strongly disagree that excerpts from ChatGPT can be copied in pre- and post-task questionnaires, Casey copied considerable amounts of ChatGPT's output into their text, as detailed in Table 4. During

²⁰ It is important to mention that participants did not receive instructions whether they could or not browse the internet for research or check information. So, even though they were instructed to use ChatGPT to improve their text, they could have checked the fake reference, but they did not. The limitations of using intentionally vague, open-ended instructions designed to simulate real-life engagement are further discussed in section 5.2 - Limitations of the study.

their performance, Casey rephrased and reworded some excerpts, as highlighted in orange in Table 4. In a similar way, in another excerpt, they adapted the last words of a ChatGPT-generated sentence. They adapted the original output from “In such a context, widespread and unrestricted use of innovations like AI **is questionable**” (emphasis added to ChatGPT’s output, Datasheets, Appendix G) to “In such a context, widespread and unrestricted use of innovations like AI **should not yet be allowed.**” (emphasis added to Casey’s Final Text, Datasheets, Appendix G).

Later, on the post-task questionnaire, Casey justified that they perceived an exaggerated use of synonyms in ChatGPT’s re-writing of the text. Because of that, they preferred to maintain their original words on certain occasions. The act of adapting a ChatGPT output that was based on their piece of writing raises attention to a relevant aspect when analyzing AI and feedback literacies: the limits and contradictions of identifying and defining authorship on AI-assisted or generated texts. To what extent is a text written by an individual, but adapted by an AI agent, truly theirs? This intriguing question has gained prominence in literature, once the boundaries between human and AI-generated text lack consensus in Academia (Teixeira da Silva; Tsigaris, 2023).

The use of ChatGPT “to write, review and edit scholarly manuscripts presents challenging ethical issues for researchers” (Hoesseini; Resnik; Holmes, 2023, p.461). First, it is challenging to measure how much of the text from a writing assignment has been AI-generated. Moreover, there is no standardized guidance on how AI-generated content should be labeled across writing guidelines. In addition to that, as LLMs are being integrated into writing workflows (such as Microsoft Word AI writing assistant²¹), questions are raised on whether it will be possible to put constraints on AI use in writing assignments. For Laquintano, Schnitzler, and Vee (2023), “the extent to which AI is constitutive to writing — or acknowledged as such — will depend on disciplinary conventions, individual writing processes, and specific implementations of the technology.” The authors point out that they see a potential for AI to be integrated into writing processes. Still, the remaining question is whether Higher Education Institutions (HEI) will be able to align with the new paradigms of AI-partnered writing or whether it will expose the artificiality of writing practices in Academia.

²¹ For more information, refer to the support article “Welcome to Copilot in Word”: <https://support.microsoft.com/en-us/office/welcome-to-copilot-in-word-2135e85f-a467-463b-b2f0-c51a46d625d1>. Accessed on 29 May 2025.

When looking back to this study, an agreement regarding authorship was set with participants: Figure 4, which presents the slideshow used during the training session, demonstrates that participants were aware that “Delegating the majority of the writing and research process to AI; [and] Generating large portions of text with little to no input from you as the author” was considered unacceptable. A question that may be raised is: how much input from the author is necessary to consider that a text is primarily human-written? Even though Casey copied most of the text from ChatGPT, the text generated was based on their input, which was original and authored by them. Even when (apparent) clear agreements are set, the boundaries of AI use and its impact on authorship are still vague and subjective, and the lack of consensus across HEIs, professors, and journals reveals a challenge which has not been regulated or normatized (Lund; Naheem, 2023).

Moreover, unlike Casey’s behavior in the feedback task, Charlie did not incorporate most of ChatGPT’s feedback. In the post-task questionnaire, they report negative feelings towards AI: “when it [ChatGPT] improved something, it worsened another aspect that I had already asked to change before. For example, at no moment, I have asked to change words and expressions, and even though I explicitly asked it not to change, it kept changing. And when I asked it to use a less pompous language, it wrote ‘like dummy’ [sic]” (Charlie’s response to post-task questionnaire, Datasheets, Appendix G). They also point out that it was a stressful experience to use ChatGPT, as they were only used to using it for brainstorming. Charlie says that ChatGPT is not useful “even for correcting text, it only confuses more, and it changes things which it should not change” (Charlie’s response to post-task questionnaire, Datasheets, Appendix G).

Negative feelings towards ChatGPT use in academic tasks have been reported in other studies. In an empirical study conducted in a Jordanian university, 20% of the participants reported feeling apprehensive due to the lack of precision in ChatGPT data and uncomfortable when using the AI chatbot (Ajlouni; Abd-Alkareem Wahba; Salem Almahaireh, 2023). Even though negative perceptions are minor in this study and in the literature, they “can result in disengagement, reduced motivation, and hindered academic success” (Tossel *et al.*, 2024, p.1070), once they are crucial when navigating learning, especially in L2 contexts (Richards, 2020). Therefore, human-AI and L2 pedagogies should also bear in mind students’ subjectivity, such as their emotions, and their perceptions and feelings towards AI, not as a limitation, but as a

variable that can inform pedagogical practices to enrich discussions around human-AI interactions.

Another problematic aspect of group C's performance on the feedback task is the use of a fake bibliographic citation derived from an AI hallucination (Walters; Wilder, 2023; Sun *et al.*, 2024). When Charlie requested during their interaction with ChatGPT to adapt the text for an academic register, the output created two fake references. On the following prompt, Charlie affirms: "I like the sources. Please maintain it, but use a simpler vocabulary [sic]". Even though participants were advised during the training session on the need for fact-checking ChatGPT outputs, mainly due to the occurrence of hallucinations (Athaluri *et al.*, 2023), Charlie incorporates the sources in their text and submits it without independently checking the references:

[...] Evidence suggests that inadequate compensation and working conditions among academic and administrative staff not only generates less satisfied professionals, but also generates declines in educational quality (Jones & Brown, 2018) [...] (Charlie's final text, Datasheets, Appendix G).

In the pre-task questionnaire, when asked about how professors should address ChatGPT in their classes, Charlie reports the lack of trustworthy academic sources as a key concern. They suggest that professors should alert students to, among other issues, the lack of [bibliographic] references. Even with such concerns, the fact that Charlie inserted a reference derived from AI hallucinations raises attention to how the apparent lack of AI literacy, in this case, their low level of AI hallucination monitoring, can be dangerous. Charlie was aware of how ChatGPT may lack bibliographic references that back up its outputs, but their level of critical monitoring of ChatGPT's output at that moment was not sufficient to avoid AI's authority and its power to convince users (Coeckelbergh, 2020; Crawford, 2023).

Once again, I emphasize that AI literacy is not a "milestone" that students reach and accomplish — even the most AI-literate individuals are subject to being deceived by GenAI due to the proximity of its output to original and reliable academic excerpts. Being AI literate is a lifelong process constantly in development, as AI technology evolves (UNESCO, 2024) and presents new paradigms. The aspect of continuing AI education will be further discussed in section 5.1, Pedagogical Implications.

Group C's behavior raises an important point when it comes to GenAI chatbots and LLMs, which is authorship. The current discussion of whether ChatGPT can be considered or not as an author when used in writing is divided into three main perspectives: liberalists, which consider ChatGPT can have authorship granted; conservationists, which deny ChatGPT authorship; and moderationists, which consider the existence of ChatGPT authorship, being aware of its lack of mental states, such as knowledge or intention, which are unique to humans (Woudenberg; Ranalli; Bracker, 2024).

The different ways in which participants approach ChatGPT authorship — whether it can be considered an author or not — are also a variable that might affect their behavior when interacting and appropriating ChatGPT's output. This will depend on the user's moral terms and to what extent they feel accountable for AI decisions (Martin, 2019). In the case of Casey, even though the submitted version of their essay is mostly AI-generated, with minor changes made by the author, the input for ChatGPT's output was written by Casey herself. On the other hand, in the case of Charlie, it is hard to argue that their text is AI-written when only a few sentences have been used from ChatGPT's suggestions, while most of the writing process was in charge of the author.

This “difficult to measure” boundary between human-machine authorship is one of the reasons why this thesis posits itself against the total banishment of AI in writing and instructional settings, a vision supported by other authors (Laquintano; Schnitzler; Vee, 2024). There are ethical precautions to bear in mind when using AI in writing (as Figure 5 exemplifies). However, LLMs are free of charge for users, and if learners can use them, they should be AI-literate enough to make the right decisions and maintain authorship of their texts (UNESCO, 2023a; UNESCO, 2023b; Kajiwarra; Kawabata, 2024).

Another aspect of the pre- and post-task questionnaire elucidated nuances in one of the participants' perceptions of AI literacy. When Casey was prompted to answer how and when professors should approach ChatGPT in their classes in the post-task questionnaire, they wrote:

Being, at the moment, the Artificial Intelligences still in development and the information society still trying to comprehend this new tool, I think it is right that universities debate its uses, but it [ChatGPT] should not be used in basic

education schools due to the lack of knowledge and critical thought from teachers and students²² (Casey, Datasheets, Appendix G).

The idea that GenAI and ChatGPT should not be approached in basic education schools is not only inadequate but also prejudiced: it reinforces the notion that knowledge and science are only produced within Academia, ignoring that schools also produce science (Ferreira Arantes; Peres, 2015). This argument is reinforced by Casey - from group C - in the post-task questionnaire when they say that basic education teachers are not ready to comprehend, talk about, or accept the use of ChatGPT in assignments.

The concept of literacy has been broadened in the last decades, integrating multiple skills and domains of knowledge that seek to empower individuals' capability of critically engaging with society and labor (The New London Group, 1996). As mentioned in Group B's pre-task questionnaire analysis (subsection 4.3.2), AI literacy is not a milestone to be reached and then "checked" from the checklist of "literacies to be developed". Considering the discussions brought within this thesis (for reference, check section 2.6), being AI literate does not only mean comprehending what AI is and how it works, but also understanding how such technology takes place in different social contexts (Coeckelberg, 2020) — including educational ones.

Finally, schools are responsible for a significant part of one's educational life. And, according to the reports I have heard from teachers and students in basic education, in many schools, students are using ChatGPT, whether they are being instructed about it or not. Hopefully, these students will reach the University level, where they will continue to have access to GenAI, LLMs, and other types of AI tools we will be researching in the following decades. It would be fantastic if those students could get to university with previous knowledge of AI and its contradictions developed at school age — the period to get a university degree may not be sufficient to explain AI from scratch.

²² Translated from the original in Portuguese: "Estando, atualmente, as Inteligências Artificiais ainda em desenvolvimento e a sociedade da informação ainda tentando compreender esta ferramenta nova, acho correto queo ambiente universitário debata seus usos, mas não que seja usado em escolas de ensino básico, por falta de conhecimento e pensamento crítico de professores e alunos" (sic).

5. CONCLUSION

This study investigated the employment and perceptions of 7 Brazilian undergraduate students on using ChatGPT as a feedback tool for an L2 writing task. In this context, the first research question elicited was “How do Brazilian undergraduate students incorporate ChatGPT’s feedback into an L2 writing assignment?”. As result of the interventions carried out in this study, participants incorporated ChatGPT feedback in three different ways: A - integral incorporation: participants integrally copied ChatGPT’s output and submitted it as their text; B - writing assistance: participants used specific suggestions and structures given by ChatGPT to edit their texts; and C - partial incorporation: participants incorporated selected excerpts of ChatGPT into their texts.

The second research question elicited was “How do Brazilian undergraduate students’ perceptions regarding ChatGPT as an AWE tool in an L2 assignment change after they engage with the feedback task?”. Alex and Ash, participants from group A, maintained their strong disagreement with the affirmation that it is acceptable to copy parts of the answers given by ChatGPT without referencing the source in both the pre-task and post-task questionnaires. While Alex increased their belief in using ChatGPT in academic tasks, Ash maintained their perception. In the post-task questionnaire, both of them agreed that ChatGPT can be used to collaborate on an academic task, to some extent. However, concerning the personal importance of understanding ChatGPT and other Artificial Intelligence (AI) tools, their views remained unchanged, though they differed between them: Alex strongly agreed that this understanding was important, whereas Ash disagreed.

Even though they strongly disagree with copying without referencing ChatGPT, group A participants engaged in integral incorporation of ChatGPT’s output into their final texts. This behavior contradicted the instructions given during the training session, which explicitly indicated that transferring most of the writing process to AI with minor author input was an unacceptable practice (refer to Figure 4). While their reported core beliefs on plagiarism did not change from pre- to post-task questionnaires, Alex and Ash demonstrated more thorough conceptualizations of ChatGPT and AI tools after the training session, when compared to their initial descriptions of the system.

This research discussed the inconsistencies between their reported beliefs and their behaviors and possible variables that could have influenced them. Among these factors, the attempt to underreport socially undesirable beliefs and to overreport socially desirable ones, also known as social desirability bias (Krumpal, 2011), is one of the possible motivations for the discrepancy between group A's performance and their answers in the questionnaires. Furthermore, other learning aspects, such as their limited prior experience with ChatGPT and potential lower levels of AI literacy, may also have led to the incongruences between their discourses and attitudes. Finally, the instructions given during data collection might not have been sufficiently clear or scaffolded. These findings underscore the complex relationships between reported perceptions, actual practice, prior ChatGPT/AI familiarity, and the potential impact of the lack of scaffolding in academic writing practices mediated by AI.

Group B, characterized by their "writing assistance" approach, exhibited more nuanced shifts in perception regarding ChatGPT as an AWE tool in L2 writing tasks. Participants from group B expressed a more flexible view towards the acceptability of incorporating ChatGPT's output into writing assignments without referencing it, yet still disagreeing with this practice. Blake, who initially strongly disagreed with this practice, later only disagreed. Bobby shifted from disagreeing to partially disagreeing, and Bennie moved from strongly disagreeing to partially disagreeing. When it comes to the belief that ChatGPT can be used to partially collaborate on an academic task, Blake shifted from agreeing to strongly agreeing, while Bobby and Bennie moved from agreeing to partially agreeing. As for the personal importance of understanding ChatGPT and AI tools, Blake moved from partially disagreeing to partially agreeing, while Bobby moved from partially agreeing to agreeing. Bennie, however, demonstrated decreased interest, shifting from partially agreeing to partially disagreeing.

Blake, Bobby, and Bennie demonstrated a more critical engagement with ChatGPT feedback. They did not incorporate ChatGPT's output integrally, as group A did, but selectively applied its suggestions to refine their texts. This behavior has been described in the literature, as other studies suggest that students with higher AI literacy skills can improve their critical thinking when interacting with GenAI (Shen and Chen, 2025). Bennie, Blake, and Bobby appear to have a higher level of AI literacy, evidenced by greater prior experience with ChatGPT and detailed conceptualization of the chatbot in questionnaires, which might have influenced their choices on the feedback task,

which led to a nuanced and selective appropriation of ChatGPT's output. In addition, all three participants advocated for open and honest discussions about ChatGPT in academic settings. This suggests an awareness of AI tools' limitations, combined with a belief in a fruitful use of such technology when properly applied.

Group C participants also decreased their level of confidence in the potential benefits of using ChatGPT to partially collaborate in writing an academic text. Charlie went from strongly agreeing to partially agreeing, while Casey went from partial disagreement to disagreement. In relation to authorship issues, Charlie maintained their partial agreement that it is acceptable to copy ChatGPT outputs without referencing them, while Casey maintained their strong disagreement with this affirmation. While Charlie affirmed that they still strongly agree that understanding how ChatGPT works is personally important for them, Casey displayed a change in their position, going from partial disagreement (in the pre-task questionnaire) to partial agreement regarding this affirmation. Finally, Charlie still maintains their argument that professors should always highlight issues such as plagiarism, originality, and critical thought at the beginning of their courses, while Casey remains arguing that this discussion is valid in Universities, but should not take place in basic education schools.

In summary, group C demonstrated decreased confidence in ChatGPT's helpfulness for academic tasks. While Casey slightly increased their interest in understanding AI, Charlie maintained their high interest, but also expressed their negative experience with using ChatGPT on the feedback task. Group C performance highlights the challenges of defining authorship (Teixeira da Silva, 2023), the risk of AI hallucinations (Sun *et al.*, 2024), and misconceptions regarding AI literacy, especially when it comes to schools, as discussed in section 4.5.2. These findings reinforce the need for comprehensive, ongoing AI education that addresses technological skills, but also critical and ethical engagement during human-AI interactions (UNESCO, 2024).

This study's findings do not indicate a single, unified answer to research question 2, which investigated how Brazilian undergraduate students' perceptions of ChatGPT as an AWE tool changed after engaging in an L2 feedback task. Instead, changes in perception were different across the three groups, and shaped by a range of interrelated factors, such as (but not limited to) participants' prior knowledge on AI and ChatGPT, the impact of the training session, and, notably, participants' experiences during the feedback task. These variables seem to have contributed to different outcomes across the three participant groups. In group B, Bobby and Blake's

growing interest in understanding AI contrasted with Bennie's decreasing interest, suggesting individual differences within the group. Group C experiences during the feedback task led to increased skepticism towards ChatGPT, as Charlie and Casey reduced their confidence regarding ChatGPT while citing issues with coherence, language, and accuracy. Group A, despite expressing strong disapproval of copying ChatGPT output, fully adopted its suggestions, revealing a dissonance between stated beliefs and their behavior, probably influenced by social desirability bias and limited AI literacy, for instance. The next subsection addresses how such findings inform pedagogical implications for the fields of L2 and literacy development.

5.1 PEDAGOGICAL IMPLICATIONS

The analysis conducted in this study brings some relevant pedagogical implications to the field of academic literacies and artificial intelligence in language education. First, it highlights the foremost relevance of debating AI in instructional settings. As AI becomes increasingly embedded in everyday life, it risks deepening existing inequities (Crawford, 2023), particularly between those who are AI literate and those who are not (UNESCO, 2019b). Providing training in public education institutions, as modeled in this study, is essential to mitigate disparities between students who can afford AI-related education and those who cannot (UNESCO, 2023a). Moreover, AI literacy must be understood as a matter of democratic participation and citizenship. With the proliferation of AI-generated content, such as deep fakes²³ or even hallucinated data, misinformation may grow (Garimella; Chauchard, 2024). As demonstrated by Charlie's incorporation of an AI-generated hallucination, even those in formal educational settings may reproduce false information, which reinforces the need to prepare society to critically engage with online content.

However, it is not advisable to assign the responsibility of turning students into AI-literate individuals to a single teacher, professor, program, or institution. This is a shared objective that should be reinforced in pedagogical projects, educational

²³ For more information on deep fakes, check <https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them>. Accessed on 26 Dec. 2024.

policies, and governmental initiatives (UNESCO, 2021). AI intersects with various fields, such as languages, technology, sciences, and beyond. For that reason, educational systems must create collaborative frameworks to equip students with the tools to use AI technologies ethically and critically, preventing their harms. Thus, avoiding talking, teaching, discussing, and studying about ChatGPT and GenAI is probably the worst path to go, as it neglects the possibility of questioning its potentialities, limitations, and pitfalls, and even worse: it does not offer individuals the chance to make an informed choice when writing a prompt or using such tools.

This study also underscores the need to address the gap between beliefs and behavior in AI use, as demonstrated by group A. Participants who rejected unethical AI use still integrally incorporated ChatGPT's output, possibly due to social desirability bias or low AI literacy (section 4.5.1). Meanwhile, group B's more critical engagement and group C's reported frustrations illustrate the relevance of scaffolding not only technical skills, but also emotional responses, ethical usage, and authorship issues. Finally, this thesis calls out the need to address the gap in human-AI interaction research in Linguistics, Literary Studies, Translation, and related areas to engage with GenAI research through diverse theoretical lenses and contribute to an ongoing, interdisciplinary conversation. There are many aspects to be investigated under a range of possible theories and approaches, as this work aimed to tie some approaches to design a method based on previous research and approaches in L2 acquisition and literacy studies (Gass; Mackey, 2015; The New London Group, 1996; Sutton, 2012). Even though limitations were encountered, this study offers a starting point for further exploration, to be addressed in the next subsections.

5.2 LIMITATIONS OF THE STUDY

While offering relevant insights into students' interactions with ChatGPT during an L2 feedback task, the development and execution of this study faced several limitations, which highlight areas of improvement for further research in this field. The first limitation relates to the emergent nature of research in AI and Applied Linguistics when the study began, in March 2023, months after ChatGPT's initial release. At the time, the field of Artificial Intelligence in Education (AIED), particularly its applications in Linguistics and Language Studies, was starting to flourish. As a result, there was a

scarce number of resources and bibliographical references to back up the study design and grounding. Many of the sources cited in this thesis, especially policy documents from UNESCO (2023a, 2023b), were published simultaneously with the research process, as they served as a starting point for debating AI effects on education from a worldwide perspective.

As empirical studies on the use of AI chatbots were scarce, this research employed an exploratory experimental design, based on pre- and post-task questionnaires, a training session, and a feedback task. The training session was based on Yan's (2023) exploratory investigation of students' behaviors when interacting with ChatGPT. In their study, students' engagement with the tool happened throughout three different phases: familiarization (learning how ChatGPT works), experimentation (practicing with ChatGPT), and exploration (hands-on work with ChatGPT).

Another methodological limitation concerns the instructions participants received when using ChatGPT. They were guided to "request feedback" to ChatGPT, an intentionally vague, open-ended prompt designed to simulate real-life engagement. However, as Tossel *et al.* (2024, pp. 1076-1077) point out,

Instructors should further experiment with integrating ChatGPT intentionally versus open-ended use by considering the iterative and dynamic nature of the writing process, as well as the specific functions that need to be performed to accomplish a writing assignment.

Therefore, an intentional use of ChatGPT, with clearer guidance on which type of prompt to work with and what to do with the output generated, could benefit not only writing as a process (Keh, 1990) but also the development of AI skills in the long run. In future studies, testing more instructional approaches (while allowing some flexibility) may help assess the pedagogical impact of intentional and guided ChatGPT use.

The absence of a control group also limited the depth of analysis regarding the training's impact. Unfortunately, it was not possible to compare the effects of different instructional conditions (e.g., trained vs. untrained, open-ended vs. intentional use), which would have allowed a more precise understanding of the training efficacy and an increase in the study's external validity. Moreover, time was also another constraint that affected the development of the study. As already mentioned, a one-and-a-half-hour training is not enough to turn participants into AI-literate individuals. AI literacy,

as this study also suggests, is not a fixed skill set but a lifelong learning process that must adapt to ongoing technological shifts. The limited timeframe restricted deeper analyses of how students' AI-related skills might evolve over longer periods of instruction and use.

Another important limitation involves individual differences and emotional factors. Although this research focused on L2 writing and literacy development, emotions and other subjective aspects of L2 development emerged during data analysis. Charlie's irritation towards ChatGPT feedback serves as an example of how emotions may have influenced participants' performance and perceptions — their prompts did not help them to achieve their writing objectives, which led to complaints and reported frustration.

Emotions, among other individual differences in L2 development, is a variable that could inform how GenAI chatbots affect students' cognition at the subjective and individual levels. These variables could have provided a deeper understanding of choices and beliefs on the feedback task and in both questionnaires, especially with participants in group A, where attitudes and beliefs strongly contrasted. Finally, other areas of literacy correlated to AI literacy were not covered. Because of this, it was not possible to understand at which point participants' lack of feedback literacy, for instance, has influenced poor choices when receiving feedback from ChatGPT. Academic and digital literacy skills could also be measured better to understand participants' performance in the feedback task.

In a general sense, these limitations underscore the complexity of students' engagement with GenAI tools and the need for nuanced, longitudinal and interdisciplinary research in human-AI interaction in learning contexts. The next section details suggestions for further research in the field based on the experiences from this thesis.

5.3 SUGGESTIONS FOR FURTHER RESEARCH

Building on the insights gained from this exploratory study and acknowledging its limitations, several gaps for further research are suggested to deepen the understanding of GenAI in L2 development settings. First and foremost, future studies should focus on collecting empirical evidence on how participants interact with LLMs

and how their interaction is related to specific variables — age, level of instruction, previous knowledge, emotions, individual differences in L2 development, literacy skills, feedback skills, digital skills, AI skills, among others.

These studies should prioritize longitudinal research designs to observe how AI literacy is developed over time. As this study and literature (UNESCO, 2024) suggest, AI literacy is not a fixed skill but a continuous and evolving process that must accompany the advancement of technology. Such longitudinal approaches would benefit from a comprehensive understanding of participants' profiles, as mentioned above. Mainly, learners' emotions also emerged in this study as an influential variable in shaping behaviors and attitudes towards ChatGPT. Future studies should integrate affective variables to better understand the subjective dimension of human-AI interaction in learning contexts.

Additionally, more experimental studies with control groups are needed, as they enable researchers to understand the extent to which their interventions affected participants' behaviors. Isolating the effects of specific interventions (such as AI literacy training or prompt writing strategies) can provide clearer insights into how training transforms learners' usage of the tools. Comparative designs could examine differences between participants who received structured guidance versus those exposed to open-ended use of AI chatbots, as well as the impact of training on learners' ability to increase their critical skills towards AI.

Another area of further research is the investigation of prompt writing in L2 settings. Understanding cognitive and linguistic processes in formulating prompts can inform more effective pedagogical strategies that support AI literacy development. Studies could explore how different types of prompts (such as requests for direct correction, as demonstrated in groups A and C versus the request for structured feedback from group B), underscoring how writing those queries informs their metacognitive strategies when writing in L2. Furthermore, research should address teaching interventions that promote intentional prompt use and critical assessment of AI content, fostering AI literacy and language development. This is also related to the incorporation of hallucinated references and the debate with students on questions regarding authorship and academic ethics.

Finally, researchers may also address a broader understanding of how AI literacy can be employed across educational levels, especially in basic education. In this study, Casey, a Librarianship undergraduate student, expressed their skepticism

regarding discussing/giving instruction about ChatGPT and AI at the school level, due to a perceived lack of knowledge and critical thinking among students and teachers from that context. However, introducing foundational AI literacy skills earlier could help prepare students for future academic and professional contexts where the use of GenAI may be needed. Therefore, partnerships between schools and universities towards exploring best practices when employing AI literacy activities in schools can collaborate to achieve inclusive and holistic AI development.

In conclusion, advancing research on GenAI in L2 educational contexts requires a multifaceted approach that addresses not only technological or linguistic aspects but also cognitive, emotional, ethical, and institutional factors. By combining longitudinal studies, experimental designs with control groups, and interdisciplinary approaches, further research can build a more comprehensive panorama of how students interact with AI tools, such as ChatGPT. Exploring AI use is essential to ensure that AI literacy becomes an inclusive and ethically grounded skill, while contributing to a more equitable, critical, and fair use of AI in society.

REFERENCES

- AJLOUNI, Aseel O.; ABD-ALKAREEM WAHBA, Fatima; SALEM ALMAHAIREH, Abdallah. Students' Attitudes Towards Using ChatGPT as a Learning Tool: The Case of the University of Jordan. *International Journal of Interactive Mobile Technologies (IJIM)*, [S. l.], v. 17, n. 18, p. pp. 99-117, 2023. DOI: 10.3991/ijim.v17i18.41753. Available at: <https://online-journals.org/index.php/i-jim/article/view/41753>. Accessed on 29 May 2025.
- ATHALURI, S. A. et al. Exploring the boundaries of reality: Investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. **Cureus**, v. 15, n. 4, e37432, 11 abr. 2023. DOI: 10.7759/cureus.37432. PMID: 37182055; PMCID: PMC10173677.
- AUERBACH, E. R.; WALLERSTEIN, N. **ESL for action: Teacher's guide**. New York: Prentice-Hall, 1987.
- BARROT, J.S. Trends in automated writing evaluation systems research for teaching, learning, and assessment: A bibliometric analysis. **Education and Information Technologies**, 2023. Available at: <https://doi.org/10.1007/s10639-023-12083-y>. Accessed on 5 Oct. 2023.
- BARROT, J. S. Using automated written corrective feedback in the writing classrooms: effects on L2 writing accuracy. **Computer Assisted Language Learning**, v.36, n.4, p.584-607, 2021. Available at: <https://doi.org/10.1080/09588221.2021.1936071>. Accessed on 6 Oct. 2023.
- BENDER, E. M.; GEBRU, T.; MCMILLAN-MAJOR, A.; SHMITCHELL, S. On the Dangers of Stochastic Parrots. **Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency**, 2021. DOI 10.1145/3442188.3445922.
- BOUD, D.; LAWSON, R.; THOMPSON, D. G. Does student engagement in self-assessment calibrate their judgement over time? **Assessment & Evaluation in Higher Education**, v. 38, n. 8, p. 941-956, 2013. Available at: <https://doi.org/10.1080/02602938.2013.769198>. Accessed on 6 Oct. 2023.
- BUCKINGHAM, David. Defining Digital Literacy: What Do Young People Need to Know About Digital Media? In: C. Lankshear and M. Knobel (eds.), **Digital Literacies: Concepts, Policies, and Practices**, New York, NY: Peter Lang, 2008. p. 17-32.
- BURSTEIN, J.; CHODOROW, M.; LEACOCK, C. Automated Essay Evaluation: The Criterion Online Writing Service. **AI Magazine**, v. 25, n. 3, p. 27, 2004. DOI: 10.1609/aimag.v25i3.1774. Available at: <https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1774>. Accessed on 14 May. 2025.

CARLESS, D.; BOUD, D. The development of student feedback literacy: enabling uptake of feedback. **Assessment and Evaluation in Higher Education**, v.43, n.8, p.1315-1325, 2018. Available at: <https://doi.org/10.1080/02602938.2018.1463354>. Accessed on 06 Oct. 2023.

CHANG, C.; CUNNINGHAM, K. J.; SATAR, H. M.; STROBL, C. Electronic feedback on second language writing: A retrospective and prospective essay on multimodality. **Writing and Pedagogy**, v.9, n.3, pp.405-428, 2018. Available at: <https://doi.org/10.1558/wap.32515>. Accessed on 13 Jun. 2024.

CHUNG, J.; MONROE, G. S. Exploring social desirability bias. **Journal of Business Ethics**, v. 44, p. 291-302, 2003. Available at: <https://doi.org/10.1023/A:1023648703356>. Accessed on 12 May 2025.

COECKELBERGH, M. **AI ethics**. Cambridge, MA: The MIT Press, 2020. 229 p.

CONSELHO NACIONAL DE SAÚDE. Resolução n o 466, de 12 de dezembro de 2012. Brasília, 2012. Available at: <https://conselho.saude.gov.br/resolucoes/2012/Reso466.pdf>. Accessed on 13 Jun. 2024.

COTTON, D. R. E.; COTTON, P. A.; SHIPWAY, J. R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. **Innovations in Education and Teaching International**, 2023. Available at: <https://doi.org/10.1080/14703297.2023.2190148>. Accessed on 6 Oct. 2023.

COUNCIL OF EUROPE. Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. 2024. Available at: <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>. Accessed on 6 Oct. 2023.

CRAWFORD, K. **Atlas de Inteligencia Artificial: Poder, Política y Costos Planetarios**. 1 ed. Buenos Aires: Fondo de Cultura Económica, 2023.

DÖRNYEI, Z. **Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Method Methodologies**. Oxford University Press: Oxford, 2011.

FAN, N; MA, Y. The Effects of Automated Writing Evaluation (AWE) Feedback on Students' English Writing Quality: A Systematic Literature Review. **Language Teaching Research Quarterly**, v. 28, p.53-73, 2022. <https://doi.org/10.32038/ltrq.2022.28.03>. Accessed on 5 Oct. 2023.

FERREIRA ARANTES, S. de L.; PERES, S. O. Programas de iniciação científica para o ensino médio no Brasil: educação científica e inclusão social. **Revista Pesquisas e Práticas Psicossociais**, v. 10, n. 1, p. 37-54, 2016. Available at: http://www.seer.ufsj.edu.br/revista_ppp/article/view/Ferreira%20Arantes%2C%20Peres. Accessed on 10 May 2025.

FERRIS, D. The Case for Grammar Correction in L2 Writing Classes: A Response to Truscott (1996). **Journal of Second Language Writing**, Vol. 8, N.1, pp.1-11. 1999.

FLECKENSTEIN, J. LIEBENOW, Lucas W.; MEYER, Jennifer. Automated feedback and writing: a multi-level meta-analysis of effects on students' performance.

Frontiers in Artificial Intelligence, v.6, 2023.

<https://doi.org/10.3389/frai.2023.1162454>. Accessed on 5 Oct. 2023.

FLOWER, L.; HAYES, J. R. A Cognitive Process Theory of Writing. **College Composition and Communication**, v.32, n. 4, 1981. DOI: 10.2307/356600.

GARCÍA-PEÑALVO, F. J.; VÁZQUEZ-INGELMO, A. What Do We Mean by GenAI? A Systematic Mapping of The Evolution, Trends, and Techniques Involved in Generative AI. **International Journal of Interactive Multimedia and Artificial Intelligence**, v. 8, n. 4, p. 7-16. Available at: doi.org/10.9781/ijimai.2023.07.006. Accessed on 16 Sep. 2025.

GARIMELLA, K.; CHAUCHARD, S. How prevalent is AI misinformation? What our studies in India show so far. **Nature**, v. 630, n. 8015, p. 32-34, 2024. Available at: <https://doi.org/10.1038/d41586-024-01588-2>. Accessed on 3 Jul. 2025.

GASS, S.M.; MACKEY, A. Input, Interaction, and Output in Second Language Acquisition. In: VANPATTEN, Bill; WILLIAMS, Jessica. **Theories in Second Language Acquisition: An Introduction**. Routledge: New York and Oxon, 2015. ISBN: 978-0-203-62894-2.

GEE, J. P. The New Literacy Studies. In: ROSWELL, J; PAHL, K. (Orgs.) **The Routledge Handbook of Literacy Studies**. Routledge: Oxon and New York, 2015. ISBN: 978-1-3157-1764-7.

HATTIE, J.; TIMPERLEY, H. The power of feedback. **Review of Educational Research**, v. 77, n. 1. 2007. pp.81-112. DOI: 10.3102/003465430298487.

HOLMES, W.; PERSSON, J.; CHOUNTA, I.A.; WASSON, B; DIMITROVA, V. **Artificial Intelligence and Education: A critical view through the lens of human rights, democracy, and the rule of law**. Council of Europe Publishing: Strasbourg, November 2022. Available at: <https://rm.coe.int/artificial-intelligence-and-education-a-critical-view-through-the-lens/1680a886bd>. Accessed on 06 Oct. 2023.

HOSSEINI, Mohammad; RESNIK, David B.; HOLMES, Kristi. The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. **Research Ethics**, v. 19, n. 4, p. 449-465, 2023. Available at: <https://doi.org/10.1177/17470161231180449>. Accessed on 27 May 2025.

HU, C.; CAO, R.; HUANG, J.; WEI, Y. The effect of self-discrepancy on online behavior: a literature review. **Frontiers in Psychology**, v. 13, 25 abr. 2022. Seção: Human-Media Interaction. Available at: <https://doi.org/10.3389/fpsyg.2022.883736>. Accessed on 9 May. 2025.

HYLAND, K; HYLAND, F. Feedback on second language students' writing. **Language Teaching**, v. 39, n. 2, p. 83-101, 2006. Available at: <https://doi.org/10.1017/S0261444806003399>. Accessed on 05 Oct. 2023.

HYLAND, K. Genre pedagogy: Language, literacy and L2 writing instruction. **Journal of Second Language Writing**, v. 16, p. 148-164, 2017. Available at: <https://doi.org/10.1016/j.jslw.2007.07.005>. Accessed on 8 Jun. 2024.

KAJIWARA, Y.; KAWABATA, K. AI literacy for ethical use of chatbot: Will students accept AI ethics? **Computers and Education: Artificial Intelligence**, v. 6, jun. 2024. Available at: <https://doi.org/10.1016/j.caeai.2024.100251>. Accessed on 10 May 2025.

KASNECI, E.; SESSLER, K.; KÜCHEMANN, S.; BANNERT, M.; DEMENTIEVA, D.; FISCHER, F.; GASSER, U.; GROH, G.; GÜNNEMANN, S.; HÜLLERMEIER, E.; KRUSCHE, S.; KUTYNIOK, G.; MICHAELI, T.; NERDEL, C.; PFEFFER, J.; POQUET, O.; SAILER, M.; SCHMIDT, A.; SEIDEL, T.; KASNECI, G. ChatGPT for good? On opportunities and challenges of large language models for education. **Learning and Individual Differences**, v. 103. 2023. Available at: <https://doi.org/10.1016/j.lindif.2023.102274>. Accessed on 05 Oct. 2023.

KEH, C. L. Feedback in the writing process: A model and methods for implementation. **ELT Journal**, v. 44, n. 4, p. 294-304. 1990. DOI:10.1093/elt/44.4.294

KOLTOVSKAIA, S. Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. **Assessing Writing**, v. 44, 2020. Available at: <https://doi.org/10.1016/j.asw.2020.100450>. Accessed on 15 Oct. 2024.

KRASHEN, S. **The Input Hypothesis: issues and implications**. Harlow: Longman, 1985.

KRUMPAL, I. Determinants of social desirability bias in sensitive surveys: a literature review. **Quality & Quantity**, v. 47, n. 4, p. 2025-2047, 2011. Available at: <https://doi.org/10.1007/s11135-011-9640-9>. Accessed on 7 May 2025.

LAQUINTANO, T.; SCHNITZLER, C.; VEE, A. Introduction to teaching with text generation technologies. In: VEE, A.; LAQUINTANO, T.; SCHNITZLER, C. (ed.). **TextGenEd: Teaching with Text Generation Technologies**. The WAC Clearinghouse, 2023. Available at: <https://doi.org/10.37514/TWR-J.2023.1.1.02>. Accessed on 14 May 2025.

LEA, M. R.; STREET, B. V. Student writing in higher education: An academic literacies approach. **Studies in Higher Education**, v.23, n.2, p.157-172. 1998.

LINK, S.; MEHRZAD, M.; RAHIMI, M. Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. **Computer Assisted Language Learning**, v. 32, n. 4, p. 605-634, 2020. Available at: <https://doi.org/10.1080/09588221.2020.1743323>. Accessed on 6 Oct. 2023.

LOEWEN, S. Uptake in Incidental Focus on Form in Meaning-Focused ESL Lessons. **Language Learning**, v. 54, p. 153-188, 2004. Available at: <https://doi.org/10.1111/j.1467-9922.2004.00251.x>. Accessed on 5 Oct. 2023.

LUND, Brady D.; NAHEEM, K. T. Can ChatGPT be an author? A study of artificial intelligence authorship policies in top academic journals. **Learned Publishing**, v. 37, n. 1, p. 13-21, 2024. Available at: <https://doi.org/10.1002/leap.1582>. Accessed on 14 May 2025.

MARTIN, K. Ethical implications and accountability of algorithms. **Journal of Business Ethics**, v. 160, p. 835-850, 2019. Available at: <https://doi.org/10.1007/s10551-018-3921-3>. Accessed on 29 May 2025.

MENG, Yuan-Yuan. Written corrective feedback: a review of studies since Truscott (1996). **Studies in Applied Linguistics and TESOL**, v. 13, n. 2, 2013. Available at: <https://doi.org/10.7916/salt.v13i2.1339>. Accessed on: 01 Jun. 2025.

NG, Davy Tsz Kit; LEUNG, Jac Ka Lok; CHU, Samuel Kai Wah; QIAO, Maggie Shen. Conceptualizing AI literacy: An exploratory review. **Computers and Education: Artificial Intelligence**, v. 2, p. 100041, 2021. Available at: <https://doi.org/10.1016/j.caeai.2021.100041>. Accessed on 03 Jun. 2025.

NIEMINEN, J.H.; CARLESS, D. Feedback literacy: a critical review of an emerging concept. **Higher Education**, v. 85, p. 1381-1400, 2023. Available at: <https://doi.org/10.1007/s10734-022-00895-9>. Accessed on 13 Jun. 2024.

NYSTRAND, M.; GREENE, S.; WIEMELT, J. Where did Composition Studies Come from? **Written Communication**, v. 10, n. 3, p. 267-333. DOI: 10.1177/0741088393010003001.

OPENAI. Introducing ChatGPT. 2022. Available at: <https://openai.com/index/chatgpt/>. Accessed on 03 Jun. 2025.

OPENAI. ChatGPT Shared Links FAQ. 2023a Available at: <https://help.openai.com/en/articles/7925741-chatgpt-shared-links-faq>. Accessed on 14 Nov. 2023.

OPENAI. How can ChatGPT be used for assessment and feedback? 2023b. Available at: <https://help.openai.com/en/articles/8313397-how-can-chatgpt-be-used-for-assessment-and-feedback>. Accessed on 10 Nov. 2023.

OPENAI. How ChatGPT and Our Language Models Are Developed. 2023c. Available at: <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed>. Accessed on 10 Nov. 2023.

OPENAI. Text corrections and feedback. ChatGPT (September 25 Version) [Large language model]. 2023d. Available at: <https://chat.openai.com/share/f1cf5936-4e27-4618-9201-2ee52c93b885>. Accessed on 06 Oct. 2023.

PANADERO, Ernesto; LIPNEVICH, Anastasiya A. A review of feedback models and

typologies: Towards an integrative model of feedback elements. **Educational Research Review**, v. 35, p. 100416, Feb. 2022. Available at: <https://doi.org/10.1016/j.edurev.2021.100416>. Accessed on 02 Jun. 2025.

RICHARDS, J. C. Exploring emotions in language teaching. **RELC Journal**, v. 53, n. 1, p. 225-239, 2020. Available at: <https://doi.org/10.1177/0033688220927531>. Accessed on 28 May 2025.

ROGERS, J; RÉVÉSZ, A. Experimental and quasi-experimental designs. In: MCKINLEY, Jim; ROSE, Heath (Eds.). **The Routledge Handbook of Research Methods in Applied Linguistics**. 1 ed. London: Routledge, 2019.

SANG, Y. Expanded Territories of “Literacy”: New Literacies and Multiliteracies. **Journal of Education and Practice**, v. 8, n. 8, 2017. ISSN 2222-288X (Online).

SCHLOSSER, Matthew. **Raising university students’ awareness of Automated Writing Evaluation tools in the English as a Foreign Language writing classroom to promote uptake of feedback and learner autonomy**. Master’s Thesis (MA in Technology-Mediated Language Teaching and Learning) – Faculty of Arts and Humanities, Universitat Oberta de Catalunya. Barcelona, p. 38, 2023. Available at: <http://hdl.handle.net/10609/148505>. Accessed on 8 Oct. 2023.

SCHMIDT, R.; FROTA, S. Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In: DAY, R. (Ed.), **Talking to learn: Conversation in second language acquisition**. Rowley, MA: Newbury, 1986.

SHEN, Y.; CHEN, L. ‘Critical chatting’ or ‘casual cheating’: how graduate EFL students utilize ChatGPT for academic writing. **Computer Assisted Language Learning**, p. 1-29, 2025. Available at: <https://doi.org/10.1080/09588221.2025.2479141>. Accessed on 7 May 2025.

SUN, Yujie; SHENG, Dongfang; ZHOU, Zihan; WU, Yifei. AI hallucination: Towards a comprehensive classification of distorted information in artificial intelligence-generated content. **Humanities and Social Sciences Communications**, v. 11, p. 1278, 2024. Available at: <https://doi.org/10.1057/s41599-024-03811-x>. Accessed on 28 May 2025.

SUTTON, P. Conceptualizing feedback literacy: knowing, being, and acting. **Innovations in Education and Teaching International**, v. 49, n. 1, p. 31-40. 2012. Available at: <https://doi.org/10.1080/14703297.2012.647781>. Accessed on 13 Jun. 2023.

SWAIN, M. Communicative competence: some roles of comprehensible input and comprehensible output in its development. In: GASS, S.; MADDEN, C. **Input in second language acquisition**. Cambridge: Newbury House, 1985, pp. 235-253.

TEIXEIRA DA SILVA, Jaime A.; TSIGARIS, Panagiotis. Human- and AI-based authorship: principles and ethics. **Learned Publishing**, v. 36, n. 3, p. 453-462, Jul. 2023. Available at: <https://doi.org/10.1002/leap.1547>. Accessed on 14 May 2025.

THE NEW LONDON GROUP. A Pedagogy of Multiliteracies: Designing Social Futures. **Harvard Educational Review**, v. 66, 1 Ed. 1996.

TOSSELL, C. C.; TENHUNDFELD, N. L.; MOMEN, A.; COOLEY, K.; DE VISSER, E. J. Student perceptions of ChatGPT use in a college essay assignment: implications for learning, grading, and trust in artificial intelligence. **IEEE Transactions on Learning Technologies**, v. 17, p. 1069-1081, 2024. DOI: <https://doi.org/10.1109/TLT.2024.3355015>. Accessed on 28 May 2025.

TRUSCOTT, J. The Case Against Grammar Correction in L2 Writing Classes. **Language Learning**, v. 46, n. 2, p. 327-369. 1996.

TRUSCOTT, John. The Case for “The Case Against Grammar Correction in L2 Writing Classes”: A Response to Ferris. **Journal of Second Language Writing**, v. 8, n. 2, p. 111-122. 1999.

UNESCO. **AI and education: guidance for policy-makers**. Paris: Unesco, 2021. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000376709.locale=en>. Accessed on 26 Sep. 2024.

UNESCO. **AI competency framework for teachers**. Paris: United Nations Educational, Scientific and Cultural Organization, 2024. ISBN 978-92-3-100707-1. Available at: <https://doi.org/10.54675/ZJTE2084>. Accessed on 29 May 2025.

UNESCO. **Beijing Consensus on Artificial Intelligence and Education**. Paris: Unesco, 2019b. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000368303>. Accessed on 06 Oct. 2023.

UNESCO. **ChatGPT and Artificial Intelligence in higher education: Quick start guide**. UNESCO and IELSAC: Paris and Caracas, 2023a. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000385146.locale=en>. Accessed on 5 Oct. 2024.

UNESCO. **Harnessing the Era of Artificial Intelligence in Higher Education: A Primer for Higher Education Stakeholders**. UNESCO and IELSAC: Paris and Caracas, 2023b. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000386670.locale=en>. Accessed on 5 Oct. 2023.

UNESCO. **International Conference on Artificial Intelligence and Education Final Report: Planning the Education in the AI Era: Lead the leap**. Paris: Unesco, 2019a. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000370967.locale=en>. Accessed on 06 Oct. 2023.

VANPATTEN, Bill; WILLIAMS, Jessica. **Theories in Second Language Acquisition: An Introduction**. Routledge: New York and Oxon, 2015. ISBN: 978-0-203-62894-2.

WALTERS, W. H.; WILDER, E. I. Fabrication and errors in the bibliographic citations generated by ChatGPT. *Sci Rep*, v. 13, n.14045. 2023. Available at: <https://doi.org/10.1038/s41598-023-41032-5>. Accessed on 24 Dec. 2024.

WANG, A. Is My Toddler a Stochastic Parrot? *The New Yorker*. 2023. Available at: <https://www.newyorker.com/humor/sketchbook/is-my-toddler-a-stochastic-parrot>. Accessed on: 26 Dec. 2024.

YAN, D. Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. **Educ Inf Technol**. 2023. Available at: <https://doi.org/10.1007/s10639-023-11742-4>. Accessed on 7 Oct. 2023.

YU, S.; LEE, I. Peer feedback in second language writing. **Language Teaching**, v. 49, n. 4. 2016. Available at: <https://doi.org/10.1017/S0261444816000161>. Accessed on 13 Jun. 2024.

ZACCARON, R. **The more the merrier (?): The impact of individual and collaborative strategic planning on performance of an oral task by young learners of English as an L2 in Brazil**. Master's Thesis (Master's degree in Linguistics) – Graduate Program in English: Linguistic and Literary Studies, Universidade Federal de Santa Catarina. Florianópolis, p. 239, 2018. Available at: <https://repositorio.ufsc.br/handle/123456789/189173>. Accessed on 16 Sep. 2025.

ZACCARON, R.; XHAF AJ, D. C. P. Knowing me, knowing you: A comparative study on the effects of anonymous and conference peer feedback on the writing of learners of English as an additional language **System**, v. 95. 2020. Available at: <https://doi.org/10.1016/j.system.2020.102367>. Accessed on 19 May 2024.

ZHANG, Z. (Victor); HYLAND, K. Fostering student engagement with feedback: An integrated approach. **Assessing Writing**, v. 51. 2022. Available at: <https://doi.org/10.1016/j.asw.2021.100586>. Accessed on 5 Oct. 2023.

ZHANG, Z. (Victor); HYLAND, K. The role of digital literacy in student engagement with automated writing evaluation (AWE) feedback on second language writing. **Computer Assisted Language Learning**, p. 1-26. 2023. Available at: <https://doi.org/10.1080/09588221.2023.2256815>. Accessed on 5 Oct. 2023.

APPENDIX A - Termo de Consentimento Livre e Esclarecido (TCLE)

Olá, você está sendo convidado/a a participar de um projeto de pesquisa sobre feedback em textos escritos em inglês, intitulado “Avaliando o uso do feedback automático do ChatGPT para a escrita em inglês por estudantes de graduação brasileiros”. Você está sendo convidado/a pois está matriculado na disciplina LLE 7409 - Tópico Especial em Língua Inglesa II, que nesse semestre terá um foco na escrita acadêmica em língua inglesa. Esse estudo está sendo conduzido pelo mestrando Vítor Pluceno Behnck (do Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários da UFSC) como sua pesquisa de mestrado, no qual é supervisionado pela Prof.^a Dr.^a Donesca Cristina Puntel Xhafaj (do Depto. de Língua e Literatura Estrangeiras da UFSC). Pedimos que você leia este formulário e tire suas dúvidas antes de concordar em participar do estudo.

Objetivo da Pesquisa:

Os objetivos deste estudo são investigar (1) a utilização do feedback escrito do ChatGPT por estudantes de graduação em suas produções escritas em língua inglesa e (2) crenças e comportamentos a respeito do uso dessa ferramenta para fins de escrita acadêmica em inglês como língua estrangeira

Procedimentos:

Você será solicitado/a a desempenhar as seguintes tarefas: (1) responder a dois questionários sobre o seu perfil de aprendiz de inglês como língua estrangeira e também sobre sua experiência com o ChatGPT para a escrita, no início e no fim da pesquisa; (2) permitir acesso a 2 ensaios (1^a e 2^a versão) em inglês escritos por você; (3) participar de uma sessão de interação com o ChatGPT para solicitar feedback ao sistema; e (5) responder a uma entrevista sobre essa experiência em dia e horário a ser combinado com cada participante – que pode acontecer presencialmente, na sala da Prof.^a Donesca, na UFSC (sala 121, CCE B), ou de forma remota, através da plataforma Google Meet, conforme a preferência do/a participante.

O tempo de dedicação esperado dos participantes é de cerca de 1h para preencher os questionários, cerca de 3h a 4h para participar da sessão de feedback, e 40min para a entrevista (o tempo necessário para preencher o formulário e fazer a entrevista final pode variar entre os participantes), totalizando 5h de dedicação à pesquisa.

Haverá algum risco na realização dessas tarefas?

As tarefas desse estudo são parecidas com outras tarefas já realizadas pelos participantes durante a graduação, por isso os riscos são mínimos; no entanto, elencamos alguns deles

abaixo, justamente com as medidas tomadas para evitar qualquer tipo de prejuízo para os participantes do estudo.

1. É possível que você sinta ansiedade ao executar uma ou mais tarefas propostas. Para diminuir as chances de isso acontecer, as instruções serão dadas da forma mais clara possível e o participante sempre poderá tirar quaisquer dúvidas que tiver com o pesquisador ou sua orientadora. Com relação à ansiedade/insegurança que pode surgir no momento de solicitar feedback para o ChatGPT, todos os participantes receberão um breve treinamento sobre como utilizar ferramentas de geração automática de texto e Inteligência Artificial.

2. Também é possível que você sinta cansaço ao executar uma ou mais tarefas propostas. Para que isso seja evitado, o questionário final ficará disponível para preenchimento ao longo de alguns dias, assim, o participante pode se engajar com essa tarefa no momento que melhor lhe convier. Com relação ao cansaço durante a entrevista final, será facultado ao participante o direito de interromper a entrevista e continuá-la em outro momento. O primeiro questionário e as duas versões do ensaio serão feitos em sala de aula, sem limite de tempo, e o participante pode fazê-los no seu tempo.

3. Outra possibilidade é que você experimente emoções/sentimentos negativos ao ler o feedback que o ChatGPT deu para o seu texto ou mesmo durante o momento que tem como objetivo a solicitação desse feedback ao próprio ChatGPT. Emoções negativas em momentos de feedback já foram relatadas em pesquisas. No entanto, ao compartilhar essas emoções com os pesquisadores - caso o participante assim deseje - surge a possibilidade do desenvolvimento do letramento para o feedback, o que vai levar o indivíduo a se beneficiar mais dessas oportunidades (momentos em que recebe feedback).

4. É importante ressaltar que a geração automática de texto por meio de modelos de linguagem pode inadvertidamente introduzir vieses nas mensagens geradas. Estes vieses podem surgir devido à natureza dos dados utilizados para treinar tais modelos, os quais refletem as tendências presentes nas bases de dados que utilizam. Para evitar o potencial impacto negativo dessas interações, todos os participantes receberão treinamento sobre como elaborar prompts para solicitar feedback construtivo.

Haverá algum benefício?

Não há previsão de nenhum benefício explícito para a participação nesse estudo. De qualquer modo, é possível que você se beneficie ao participar neste estudo no que diz respeito ao aprendizado da língua estrangeira, uma vez que é documentado que o processo de feedback leva os aprendizes a refletirem sobre a língua. É possível, ainda, que ao receber feedback do sistema, você desenvolva estratégias de revisão de texto que poderá usar no futuro. Outro possível benefício é o desenvolvimento de conhecimentos a respeito da Inteligência Artificial ao longo desse processo. Por último, é importante ressaltarmos que o estudo tem um propósito social, que é permitir o avanço das pesquisas no Brasil com relação aos processos envolvidos na aprendizagem de uma segunda língua em um contexto instrucional (isto é, quando os aprendizes fazem aula de língua). Assim, a longo prazo, espera-se que todos os aprendizes nesse contexto colham benefícios de pesquisas na área de ensino e aprendizagem de segunda língua.

A sua identidade será revelada?

Não. A fim de proteger a identidade dos participantes, pseudônimos serão utilizados nas publicações a fim de evitar sua identificação.

Haverá acompanhamento de alguém?

Sim, durante todas as etapas da pesquisa o mestrando Vítor Pluceno Behnck e a Prof.^a Donesca estarão disponíveis e quaisquer dúvidas serão esclarecidas.

A participação nessa pesquisa é obrigatória a todos os alunos cursando a disciplina?

Não. A participação é totalmente voluntária e não fará parte da grade curricular da disciplina. Esse documento é um convite. Caso haja a recusa na participação, isso não terá impacto algum na sua nota ou no seu relacionamento com a professor/a da disciplina ou com o pesquisador. Durante toda a pesquisa, você também tem o direito de não responder a qualquer das perguntas feitas no questionário/entrevista.

A participação na pesquisa acarretará alguma despesa?

Não são previstas despesas para os participantes da pesquisa, uma vez que a coleta de dados acontecerá na UFSC e a ideia é que isso ocorra em horários nos quais os participantes já estarão na UFSC em função de aulas ou outras atividades.

Haverá benefício financeiro?

Não. A participação na pesquisa é voluntária e não envolve dinheiro, mas caso você venha a sofrer qualquer prejuízo decorrente de sua participação nessa pesquisa, você será indenizado de acordo com a legislação vigente.

É possível desistir de participar ou cancelar essa autorização?

Sim. É possível cancelar a participação a qualquer momento da pesquisa, sem nenhuma necessidade de justificativa ou prejuízo para você. Isso pode ser feito através do telefone do mestrando Vítor (48) 988323707, do seu e-mail: vitorbehnck@gmail.com ou pessoalmente ou também com a Prof.^a Donesca no donescax@gmail.com ou pessoalmente.

Ademais, há a garantia de livre acesso às informações da pesquisa.

Caso você queira, você pode entrar em contato com o Comitê de Ética em Pesquisas com Seres Humanos (CEPSH) da UFSC. O CEPSH é um órgão colegiado interdisciplinar, deliberativo, consultivo e educativo, vinculado às universidades, mas independente na tomada de decisões, criado para defender os interesses dos participantes da pesquisa em sua integridade e dignidade e para contribuir no desenvolvimento da pesquisa dentro de padrões éticos. O contato com o CEPSH, na UFSC, pode ser feito pessoalmente, por telefone, ou por e-mail, utilizando os dados fornecidos a seguir:

Prédio Reitoria II, 7º andar, sala 701, localizado na Rua Desembargador Vitor Lima, nº 222, Trindade, Florianópolis/SC.

Telefone para contato: (48) 3721-6094 - Email: cep.propesq@contato.ufsc.br

Os procedimentos metodológicos adotados obedecem aos preceitos éticos da pesquisa, conforme normatizado pela resolução do CNS 510 de 2016. Os pesquisadores se comprometem a seguir tal resolução, bem como declaram conhecer e cumprir os requisitos da Lei Geral de Proteção de Dados (Lei nº. 13.709, de 14 de agosto de 2018) quanto ao tratamento de dados pessoais sensíveis.

Esse documento deverá ser assinado em duas vias, todas as páginas rubricadas, ficando uma via com você e uma com o pesquisador. Assinando o consentimento pós-informação, você

estará consentindo com o uso dos dados coletados para a pesquisa e para sua posterior divulgação, de acordo com o acima estabelecido.

Muito

obrigado,

Vítor Pluceno Behnck - Pesquisador

Donesca Cristina Puntel - Supervisora da pesquisa

Consentimento pós-informação

Eu, _____ (nome completo),
fui esclarecido/a sobre a pesquisa “Avaliando a utilização do feedback de escrita automatizada do ChatGPT por estudantes de graduação brasileiros”.

Florianópolis, ____ de _____ de 2024.

APPENDIX B - Pre-task writing test

**UNIVERSIDADE FEDERAL DE SANTA CATARINA - CENTRO DE COMUNICAÇÃO E EXPRESSÃO
DEPARTAMENTO DE LÍNGUA E LITERATURA ESTRANGEIRA**

Name: _____

Teste para diagnóstico - Escrita

- Escreva um ensaio, em língua inglesa, sobre qualquer assunto que seja do seu interesse.
- Lembre-se de: dar um título para seu texto, ter introdução (1 para.) , desenvolvimento (2 para.) , e conclusão (1 para.) .
- Na sua introdução, deixe clara qual a sua opinião sobre o assunto que você vai abordar. No desenvolvimento, traga pelo menos 2 argumentos que dão suporte à sua opinião. Na conclusão, retome sua opinião, conectando os argumentos trazidos a ela.
- Use uma linguagem mais formal.
- Cada parágrafo deve ter entre 3 e 6 linhas.

APPENDIX C - Pre-task questionnaire

Participação em pesquisa

* Indica uma pergunta obrigatória

1. E-mail * _____

Parte 1 - Dados pessoais

2. Nome * (Essa informação é confidencial e não será divulgada ou publicada)

3. Idade * _____

4. Há quanto tempo você estuda Inglês? (em anos) _____

5. Qual o seu curso de graduação? * _____

Parte 2 - Experiências prévias com ferramentas de revisão automática de texto e ChatGPT

Nessa seção você será questionado/a a respeito da sua experiência no uso de ferramentas de revisão automática de texto e o ChatGPT.

6. Você usa alguma ferramenta de correção automática de texto? *
(Ferramenta de correção do Word/Google Docs, Grammarly, dentre outros, por exemplo)

Marcar apenas uma oval.

☐ Sim ☐ Não

7. Se você respondeu "sim", escolha as ferramentas que você usa. Você pode selecionar mais de uma. *Marque todas que se aplicam.*

☐ Correção automática do Microsoft Word ☐ Correção automática do Google Docs

☐ Grammarly

☐ Dicionários Online

☐ Outro: _____

8. Com que frequência você usa o ChatGPT? * *Marcar apenas uma oval.*

- ☐ Nunca usei
☐ Usei algumas vezes para testar (menos de uma vez por mês)
☐ Uso com pouca frequência (pelo menos uma vez por mês)
☐ Uso com muita frequência (pelo menos uma vez por semana)
☐ Uso todos os dias

9. Você já usou o ChatGPT para realizar alguma tarefa em Língua Estrangeira? (Por exemplo: realizar uma tradução, escrever um email, revisar um texto) *Marcar apenas uma oval.*

☐ Sim ☐ Não

10. Caso você tenha utilizado o ChatGPT para realizar alguma tarefa em Língua Estrangeira, marque as alternativas correspondentes abaixo ou insira "Outro" em caso de a atividade não estar listada. *Marque todas que se aplicam.*

- ☐ Traduzir um texto ☐ Escrever um e-mail ☐ Revisar um texto
☐ Ajudar a estudar para uma prova ☐ Explicar um conceito ☐ Resumir um texto
☐ _____

Outro:

11. Você já usou o ChatGPT para realizar alguma atividade acadêmica? (Por exemplo: provas, trabalhos escritos, slides, dentre outros) *Marcar apenas uma oval.*

☐ Sim ☐ Não

12. Caso você tenha utilizado o ChatGPT para realizar alguma atividade acadêmica, marque as alternativas correspondentes abaixo ou insira "Outro" em caso de a atividade não estar listada. *Marque todas que se aplicam.*

- ☐ Provas escritas ☐ Trabalhos escritos ☐ Preparação de apresentações (slides)
☐ Outro: _____

13. Se você respondeu "sim" sobre o uso do ChatGPT em atividades acadêmicas na pergunta 11, responda brevemente: quais eram seus objetivos ao usar a ferramenta? Você acredita que conseguiu alcançar esses objetivos?

14. Descreva, em poucas palavras, o que é o ChatGPT e suas funcionalidades. *
Limite máximo: 300 palavras.

Parte 3 - Crenças a respeito do ChatGPT

Nessa seção você será questionado/a a respeito do seu ponto de vista a respeito do uso do ChatGPT.

Avalie as questões abaixo entre 1 a 6, sendo 1 (discordo plenamente) e 6 (concordo plenamente).

- 1 - discordo plenamente.
- 2 - discordo.
- 3 - discordo parcialmente.
- 4 - concordo parcialmente
- 5 - concordo.
- 6 - concordo plenamente.

15. O ChatGPT pode escrever textos acadêmicos de qualidade.

- 1 - disaccordo pienamente.
- 2 - disaccordo.
- 3 - disaccordo parzialmente.
- 4 - concordo parzialmente
- 5 - concordo.
- 6 - concordo pienamente.

Marcar apenas uma oval.

discordo plenamente 1 2 3 4 5 6 concordo plenamente
☐ ☐ ☐ ☐ ☐ ☐

16. O ChatGPT pode ser usado para colaborar parcialmente na escrita de um texto acadêmico (por exemplo: gerar sugestões de temas, revisar a gramática do texto, reescrever frases).

- 1- disaccordo pienamente.
- 2- disaccordo.
- 3- disaccordo parzialmente.
- 4- concordo parzialmente
- 5- concordo.
- 6- concordo pienamente.

Marcar apenas uma oval.

discredo plenamente 1 2 3 4 5 6 concordo plenamente

☐ ☐ ☐ ☐ ☐ ☐

17. É aceitável copiar partes das respostas dadas pelo ChatGPT sem referenciar a fonte.

- 1 - disaccordo pienamente.
- 2 - disaccordo.
- 3 - disaccordo parzialmente.
- 4 - concordo parzialmente
- 5 - concordo.
- 6 - concordo pienamente.

Marcar apenas uma oval.

discredo totalmente 1 2 3 4 5 6 concordo totalmente

☐ ☐ ☐ ☐ ☐ ☐

18. O ChatGPT pode ser usado como uma ferramenta de aprendizagem.

- 1- discordo plenamente.
- 2- discordo.
- 3- discordo parcialmente.
- 4- concordo parcialmente
- 5- concordo.
- 6- concordo plenamente.

Marcar apenas uma oval.

discordo plenamente 1 2 3 4 5 6 concordo plenamente

☐ ☐ ☐ ☐ ☐ ☐

19. Eu tenho facilidade em escrever comandos para o ChatGPT que atinjam meus objetivos.

- 1- discordo plenamente.
- 2- discordo.
- 3- discordo parcialmente.
- 4- concordo parcialmente
- 5- concordo.
- 6- concordo plenamente.

Marcar apenas uma oval.

discordo plenamente 1 2 3 4 5 6 concordo plenamente

☐ ☐ ☐ ☐ ☐ ☐

20. Entender como o ChatGPT e outras ferramentas de Inteligência Artificial funcionam é pessoalmente importante para mim .

- 1- discordo plenamente.
- 2- discordo.
- 3- discordo parcialmente.
- 4- concordo parcialmente
- 5- concordo.
- 6- concordo plenamente.

Marcar apenas uma oval.

discordo plenamente 1 2 3 4 5 6 concordo plenamente

☐ ☐ ☐ ☐ ☐ ☐

21. Entender como o ChatGPT e outras ferramentas de Inteligência Artificial funcionam é importante para o futuro do meu trabalho.

- 1- discordo plenamente.
- 2- discordo.
- 3- discordo parcialmente.
- 4- concordo parcialmente
- 5- concordo.
- 6- concordo plenamente.

Marcar apenas uma oval.

discordo plenamente 1 2 3 4 5 6 concordo plenamente

☐ ☐ ☐ ☐ ☐ ☐

22. Acredito que professores na universidade não consigam identificar se um texto foi escrito pelo ChatGPT.

- 1- discordo plenamente.
- 2- discordo.
- 3- discordo parcialmente.
- 4- concordo parcialmente
- 5- concordo.
- 6- concordo plenamente.

Marcar apenas uma oval.

discordo plenamente 1 2 3 4 5 6 concordo plenamente

☐ ☐ ☐ ☐ ☐ ☐

23. Eu não consigo identificar se um texto foi escrito por uma ferramenta de Inteligência Artificial

- 1- discordo plenamente.
- 2- discordo.
- 3- discordo parcialmente.
- 4- concordo parcialmente
- 5- concordo.
- 6- concordo plenamente.

Marcar apenas uma oval.

discordo plenamente 1 2 3 4 5 6 concordo plenamente

☐ ☐ ☐ ☐ ☐ ☐

24. De acordo com a sua opinião, como e quando o professor deveria conversar com os alunos sobre o ChatGPT em suas aulas? Limite máximo: 300 palavras.

APPENDIX D - Training session class plan

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS

Avaliando a utilização do feedback de escrita automatizada do ChatGPT por estudantes de graduação brasileiros

Vítor Pluceno Behnck
Orientadora: Donesca C. P. Xhafaj, Dra.

PLANO DE AULA - SESSÃO DE FEEDBACK COM O CHATGPT
Tempo total: 4h.

1. Introdução (30min)
• Introduzir a sessão mencionando algumas ferramentas de correção automática de texto (Grammarly, Corretor do Microsoft Word/Google Docs); • Explicar que essas ferramentas se baseiam na Inteligência Artificial. Introduzir o conceito de Inteligência Artificial; • Apresentar o ChatGPT, que é um chatbot que se baseia em Inteligência Artificial; • Questionar se os alunos possuem outras ideias de como utilizar o ChatGPT pra fins educacionais. Questionar quais seus limites; • Apresentar algumas possibilidades pedagógicas no seu uso, como apresentado em UNESCO (2023); • Discutir limites do ChatGPT, como plágio, por exemplo (Yan, 2023).
2. Prática com o sistema (30min)
• Solicitar aos alunos que acessem a plataforma do ChatGPT. Propor a criação de diferentes prompts relacionados a escrita e revisão de textos em inglês para que eles testem suas funcionalidades; • Perguntar se há dúvidas quanto ao funcionamento da plataforma; • Designar uma tarefa de revisão de texto para os alunos realizarem junto ao ChatGPT.
3. Sessão de feedback - coleta de dados (2h00min total, 1h30 para escrita)
• Disponibilizar aos alunos uma cópia em um documento de texto do ensaio que escreveram previamente; • Dar as seguintes instruções para o processo de solicitação de feedback: ○ Considerando as atividades e discussões realizadas até o momento, você agora irá interagir com o ChatGPT para solicitar feedback ao texto que você produziu no início do semestre. Aqui estão algumas orientações bem importantes que você deve seguir durante esse processo: ■ “Acesse o endereço exemplo.com.br para fazer download do seu texto. Abra-o no programa “Bloco de Notas”. É importante que você

não use os recursos de revisão de texto do Google Docs ou Microsoft Word, ou de qualquer outro sistema de revisão automática. ■ Considere o ChatGPT como uma ferramenta de revisão do seu texto. A partir disso, você pode e deve criar <i>prompts</i> solicitando que o sistema faça revisões no seu texto da maneira como quiser. ■ Ao finalizar, salve a nova versão do seu texto e clique na opção de compartilhamento do link do chat. ■ Por fim, depois que você tiver concluído todas as etapas, acesse o seguinte formulário “<i>forms.google.com/exemplo</i>”. Insira o link do chat que você acabou de copiar, a versão final do seu texto, e responda às perguntas do questionário”. ● Informar aos alunos a quantidade de tempo necessária para realizar o feedback. ● Ao finalizarem a sessão de feedback, disponibilizar o link do formulário pós-tarefa para que os alunos possam o preencher e inserir o link da página da sua interação com o ChatGPT, bem como a versão final do seu texto.
Recursos necessários
- Slideshow com as instruções das atividades; - Computadores do laboratório de informática do CCE/UFSC.

REFERÊNCIAS

UNESCO. **Harnessing the Era of Artificial Intelligence in Higher Education A Primer for Higher Education Stakeholders**. UNESCO and IELSAC: Paris and Caracas, 2023. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000386670.locale=en>. Acessado em: 5 out. 2023.

YAN, D. Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. **Educ Inf Technol**. 2023. Disponível em: <https://doi.org/10.1007/s10639-023-11742-4>. Acessado em: 7 out. 2023.

APPENDIX E - Instructions for task development

FEEDBACK DOS ENSAIOS

Nesse momento, nós vamos trabalhar com os ensaios elaborados no primeiro dia de aula. Você deve:

1. Acessar o link abre.ai/textosaula1, fazer download do seu texto no botão 
2. Crie um novo chat no ChatGPT, clicando no botão 
3. Entre no ChatGPT e interaja com o sistema a fim de solicitar feedback, com base nas discussões dessa oficina. Melhore o seu texto com base nessas interações;
4. Acessar o link abre.ai/atividadechatgpt2
5. Insira no formulário três coisas: 1. o link da sua interação com o ChatGPT, 2. o texto integral da sua interação com o ChatGPT, e 3. responda o questionário.
6. Quando terminar, não feche nenhuma aba nem desligue o computador.

EM CASO DE DÚVIDAS, PEÇA AJUDA

APPENDIX F - Post-task questionnaire

Formulário pós-tarefa: Avaliando o *uptake* de estudantes de graduação brasileiros da avaliação de escrita automatizada do ChatGPT

* Indica uma pergunta obrigatória

Parte 1 - Dados pessoais

1. Nome *

Essa informação é confidencial e não será divulgada ou publicada.

2. A fim de proteger sua identidade, pedimos que você crie um nome fictício para que seus dados sejam associados ao serem reportados. *

Por exemplo: João, Marta, Igor...

3. Insira aqui o link com a página da sua interação com o ChatGPT *

*Siga as instruções ou aguarde as instruções do pesquisador para preencher este campo.
Em caso de dúvidas, peça ajuda.*

4. Selecione com o mouse, copie e cole nessa área toda a sua interação com o ChatGPT (incluindo tanto as suas mensagens quanto as respostas do sistema). *

5. Insira aqui a versão final do seu texto, após o feedback do ChatGPT. *

Arquivos enviados:

Parte 2 - Autoavaliação

Nessa seção você será questionado/a sobre suas percepções a respeito da tarefa que acabou de executar.

6. Considere os comandos (*prompts*) que você escreveu para o ChatGPT. Qual tipo de prompt levou a um melhor feedback? *

7. Considere o feedback dado pelo ChatGPT. Teve algum feedback que você gostou mais? *

8. Considere o feedback dado pelo ChatGPT. Teve algum feedback que você não gostou? *

9. Você usaria o ChatGPT novamente para tarefas acadêmicas? Você faria algo de diferente? *

Parte 3 - Crenças a respeito do ChatGPT

Nessa seção você será questionado/a a respeito do seu ponto de vista a respeito do uso do ChatGPT.

Avalie as questões abaixo entre 1 a 6, sendo 1 (discordo plenamente) e 6 (concordo plenamente).

- 1 - discordo plenamente.
- 2 - discordo.
- 3 - discordo parcialmente.
- 4 - concordo parcialmente
- 5 - concordo.
- 6 - concordo plenamente.

16. Entender como o ChatGPT e outras ferramentas de Inteligência Artificial funcionam é importante para o futuro do meu trabalho.

*

- 1 - discordo pienamente.
2 - discordo.
3 - discordo parzialmente.
4 - concordo parzialmente
5 - concordo.
6 - concordo pienamente.

Marcar apenas uma oval.

1 2 3 4 5 6

disc ☐ ☐ ☒ ☐ ☐ ☐ concordo plenamente

17. Acredito que professores na universidade não consigam identificar se um texto * foi escrito pelo ChatGPT.

*

- 1 - discordo plenamente.
- 2 - discordo.
- 3 - discordo parcialmente.
- 4 - concordo parcialmente
- 5 - concordo.
- 6 - concordo plenamente.

Marcar apenas uma oval.

1 2 3 4 5 6

disc ☐ ☐ ☒ ☐ ☐ ☐ concordo plenamente

18. Eu não consigo identificar se um texto foi escrito por uma ferramenta de por *
Inteligência Artificial

- 1 - discordo plenamente.
- 2 - discordo.
- 3 - discordo parcialmente.
- 4 - concordo parcialmente
- 5 - concordo.
- 5 - concordo plenamente.

Marcar apenas uma oval.

1 2 3 4 5 6

disc ☐ ☐ ☐ ☐ ☐ ☐ concordo plenamente

19. De acordo com a sua opinião, como e quando o professor deveria conversar *
com os alunos sobre o ChatGPT em suas aulas?

Limite máximo: 300 palavras.

APPENDIX G - Data sheets

All appendixes are available in the following clickable link:
https://docs.google.com/spreadsheets/d/1mZfKdYTTTrR9GZJ5Su4nCS7a5l_dW_CybZfFAf9lY4tU/edit?usp=sharing.