# EFL Students' Perspectives on ChatGPT in Translation: Exploring Al Assistance, Motivation, and Learning Outcomes

# **Cao-Tuong Dinh**

English Language Department, FPT University, Can Tho Campus, Vietnam

TuongDC@fe.edu.vn

https://doi.org/10.34190/ejel.23.2.4006

An open access article under CC Attribution 4.0

Abstract: Recent advances in artificial intelligence (AI) offer promising opportunities to improve language education, particularly in translation, by providing tools that can enhance both learning processes and outcomes. Yet, how these AI tools are perceived and integrated, especially in areas that demand cultural sensitivity and a nuanced understanding, has not been fully explored, notably from the perspective of English as a Foreign Language (EFL) students. This study set out to examine how English major students view the use of ChatGPT, a text-based generative AI tool, within translation classes, using the Expectation-Value Theory as a framework. The study involved 62 junior English majors from a university in Vietnam and employed a mixed-methods approach, including pre- and post-course surveys, reflective journals, focus group interviews, and analysis of course grades. Results showed that students generally regarded ChatGPT as a helpful tool for improving translation accuracy, efficiency, and vocabulary skills. They valued its capacity to simplify complex translation challenges, improve sentence flow, and offer a variety of lexical choices, which in turn boosted their motivation and confidence. Students also mentioned that using ChatGPT helped promote collaborative learning by sparking more group discussions, which improved their translation skills. At the same time, they pointed out some limitations, especially how ChatGPT struggled with cultural nuances and idiomatic expressions. Because of this, students had to carefully review and adjust the Al's suggestions themselves. The study points out that it's really important to strike a balance between relying on Al tools and sticking with traditional, hands-on translation methods. Tools like ChatGPT can definitely support translation learning, but they can't take the place of human judgment and effort. When looking at why students were motivated to use ChatGPT, the research found that it mostly came down to how helpful they believed the tool was, how much they valued using it, and how confident they felt about succeeding with it. These factors played a key role in their overall learning results. The study provides useful insights into how AI tools can support online learning by making it more efficient and engaging. However, it also reminds us that human judgment remains crucial, especially in translation tasks that involve cultural understanding. More research is needed to understand the long-term effects of AI in translation education and how well these tools work across different cultural settings.

Keywords: ChatGPT, AI in translation, Motivation, Engagement, Expectancy-value theory

# 1. Introduction

In recent years, artificial intelligence (AI) has made remarkable progress, reshaping many aspects of education, especially in language learning and translation. Among the various AI tools available, large language models (LLMs) have attracted growing interest because of their ability to support learners in a range of language-related activities, including generating text, answering questions, and translating (Barrot, 2024; Karataş, Yaşar, and Gunyel, 2024). Their capacity to produce text that closely resembles human writing has made them valuable additions to language and translation teaching, offering new possibilities that were difficult to achieve with traditional approaches (Barrot, 2024; Karataş, Yaşar and Gunyel, 2024). However, despite their growing use, the integration of AI into educational contexts, particularly in translation education, remains underexplored in empirical research (Sahari, Al-Kadi and Ali, 2023; Xiao and Zhi, 2024). The rapid growth of generative AI has prompted a need for more in-depth investigation into its applications, effectiveness, and challenges, especially when deployed as a learning tool in translation tasks (Lo, et al., 2024; Salloum, et al., 2024).

Recent studies have explored the implications of AI in second language acquisition (SLA) and translation pedagogy, highlighting both its advantages and limitations. AI tools have been shown to enhance students' engagement with linguistic content, offer personalized learning experiences, and provide immediate feedback that can foster self-regulated learning (Sahari, Al-Kadi and Ali, 2023; Xiao and Zhi, 2024). These tools also serve as cognitive scaffolds, assisting learners in tackling complex linguistic structures and increasing their autonomy in the learning process (Karataş, et al., 2024; Xiao and Zhi, 2024). However, despite these advantages, there remain substantial challenges that warrant further investigation. AI translation tools perform well with sentence structure and straightforward translations. However, they frequently struggle to capture deeper meanings, idiomatic expressions, and cultural nuances, areas where human judgment remains essential to ensure accuracy

ISSN 1479-4403 99 ©The Authors

and quality (Ghassemiazghandi, 2024; Sahari, Al-Kadi and Ali, 2023; Van Horn, 2024). On top of that, the quality of Al translations can vary a lot depending on the language pair or the type of text, which makes it hard to rely on Al across all kinds of translation tasks. This shows that while Al has great potential, we also need research that looks beyond technology and considers its real-world limits, especially in teaching environments.

One major gap in current research is how students mentally engage with AI-generated translations, especially in advanced translation courses. While earlier studies have looked at AI's general usefulness for translation and language learning, few have explored how students critically interact with AI outputs or how these tools might actually help students develop deeper learning strategies in translation education. Additionally, the intersection between AI-assisted learning and self-regulated learning strategies, which are crucial for promoting autonomous learning, remains an area of limited research.

Focusing specifically on EFL learners is crucial because these students face unique linguistic and cultural challenges in translation tasks that AI tools often cannot fully address. Motivation plays a pivotal role in this context. Yet, there is limited emperical evidence on how students' motivation to use AI tools is influenced by their perceptions of these tools' utility, intrinsic value, and expectancy for success in translation activities. Understanding these motivational and cognitive dimensions is fundamental for effectively integrating AI into translation pedagogy and designing instructional approaches that promote balanced reliance on technology and manual skills. Therefore, this study aims to fill these gaps by focusing on how English major students in advanced translation courses perceive the value of AI-driven tools, using the Expectation-Value Theory (Wigfield and Eccles, 2000) as the guiding framework. This theory underscores the role of students' motivation, shaped by their perceptions of AI's utility and relevance, in determining their engagement with AI tools and the effectiveness of these tools in enhancing their translation performance.

Despite these promising applications, AI tools have notable limitations, particularly when handling complex cultural nuances and idiomatic expressions. These challenges underscore the need for further investigation into the pedagogical effectiveness of AI tools in real-world translation tasks. While AI translation models are highly effective at processing syntax and literal translations, they frequently fall short in understanding context, tone, and cultural appropriateness, which are essential for high-quality translations (McIntosh, et al., 2025; Sahari, Al-Kadi and AIi, 2023). This research goes beyond simply evaluating the technical capabilities of AI in translation; it takes a closer look at its impact on teaching and learning. In particular, it focuses on how students perceive AI's role in improving translation accuracy and how well it fits with their learning goals. The study also considers the wider effects of bringing AI into the classroom, especially how these tools affect student motivation and learning outcomes. More specifically, it aims to understand how students' views on the reliability of AI and its potential to improve translation quality influence their overall engagement with these technologies. With AI becoming increasingly common in education, this study responds to the pressing need to explore how it can be integrated effectively into translation teaching while also recognizing its limitations and the importance of students critically interacting with AI outputs.

By examining EFL students' perceptions and mental engagement with AI in translation, this research hopes to offer useful insights on how AI can support learning in translation education, without losing sight of the crucial balance between technological support and human expertise.

# 2. Literature Review

# 2.1 ChatGPT's Role in EFL Learning and Language Education

Generative AI has been increasingly recognized as an important aid in language learning, offering capabilities such as text generation, question answering, and translation (Alawida, et al., 2023). Its role in language learning is increasingly being explored, with studies indicating its potential as a knowledgeable learning companion (Solak, 2024). Students have shown the capability of critically evaluating and modifying ChatGPT's outputs, which can enhance learning experiences and offset academic integrity concerns (Xiao and Zhi, 2023). Additionally, generative AI applications has been recognized for its usefulness in informal digital learning environments, where it supports EFL learners in engaging with language tasks creatively and productively (Liu and Ma, 2023; Li, 2024).

Several empirical studies underscore the pedagogical value of generative AI in enhancing learners' confidence and task completion, particularly in language-related assignments. Xiao and Zhi's (2024) study revealed that ChatGPT can facilitate students' tasks requiring language competence. In a similar vein, Van Horn (2024) looked at how Korean university EFL students feel about generative AI and found that many students had positive opinions. They especially liked how it helped improve their language accuracy and encouraged them to learn on

their own. These findings show that generative AI has a lot of potential to make language learning better for EFL students, no matter which platform is used. In this study, Van Horn's results help us understand how EFL students might view generative AI in translation classes, particularly when it comes to how useful they find it and how it supports their independence as learners.

While generative AI has demonstrated significant value in improving language learning outcomes, further investigation is needed to understand how these tools are perceived by students in the context of advanced translation courses in Vietnam. Specifically, exploring students' engagement with AI tools, their critical evaluation of AI-generated outputs, and the impact of these tools on learner autonomy provides a clear research opportunity for future studies.

#### 2.2 Using Generative AI in Translation

For instance, Ghassemiazghandi (2024) evaluated a large language model's translation accuracy using the BLEU score and found a notable improvement over traditional machine translation tools. Such findings affirm the growing accuracy and reliability of generative AI in performing translation tasks. Furthermore, studies in varied contexts, such as the Arab world, have found that Al-based tools are particularly effective in handling the mechanical aspects of translation, providing learners with a base version that can be reviewed and improved upon by humans (Ghassemiazghandi, 2024; Sahari, Al-Kadi and Ali, 2023). In informal digital settings, generative Al helps learners interact more with the material by allowing them to revise, reflect, and learn from Al-generated content right away. However, how people accept and view generative AI in translation isn't the same for everyone. Salloum and colleagues (2024) found that acceptance varies depending on the language pair, showing that AI tools need to be evaluated based on specific languages and contexts. There are also ongoing concerns about generative AI struggling with idioms, cultural details, and style. Because of these limitations, human review and editing remain important, especially for professional translations or those that require a deep understanding of context (Ruoqi, Yuan, and Gochuico, 2023; Sahari, Al-Kadi and Ali, 2023). In this study, such concerns are central to the investigation, as student reflections and evaluations provide insight into both the strengths and the limitations of using generative AI in translation education. This body of literature supports the importance of studying learners' cognitive and motivational engagement with AI outputs, especially when integrated into instructional settings.

While generative AI tools have demonstrated clear advantages in terms of efficiency and accuracy, there remains a need to investigate how these tools influence the learning outcomes of students in the context of translation education in both developed and emerging markets like Vietnam. This research can help determine the pedagogical strategies best suited for integrating AI tools into translation curricula, as well as how AI tools influence students' cognitive and motivational engagement in translation tasks.

# 2.3 Expectation-Value Theory in Language Learning

The Expectation-Value Theory (Wigfield and Eccles, 2000) suggests that students' motivation is influenced by two key factors: how much they believe they can succeed at a task (expectancies) and how important they see the task (task values). Expectancy beliefs refer to students' confidence in their ability to complete a task successfully. This confidence is shaped by their past experiences, their self-belief, and the current situation (Wigfield, Tonks, and Klauda, 2016). When students have a strong belief that they can succeed, they tend to put in more effort and stick with the task longer. Task values are multifaceted and typically decomposed into four distinct types (Wigfield, Tonks and Klauda, 2016; Trautwein, et al., 2012):

Attainment (Importance) Value: The significance an individual places on doing well on the task, often linked to identity and personal goals (Wigfield, Tonks and Klauda, 2016). For example, a learner who views translation competence as central to their identity would place high attainment value on translation tasks.

*Intrinsic Value:* The inherent enjoyment or interest a person finds in performing the task (Harackiewicz, Smith and Priniski, 2016). This reflects how engaging or pleasurable the learner perceives the task itself to be, independent of external rewards.

*Utility Value:* The perceived usefulness of the task for achieving future goals, such as career advancement or academic success (Canning and Harackiewicz, 2015). In the context of AI-assisted translation, utility value would reflect how learners perceive generative AI tools as beneficial for their learning or professional practice.

*Cost* (not requested but often included): The perceived negative aspects associated with task engagement, such as effort, time, or anxiety (Flake, et al., 2015).

These expectancy and value components jointly influence learners' academic outcomes, including their engagement, persistence, and performance (Gaspard, et al., 2015). For example, when learners expect to succeed and value the task highly, they are more likely to dedicate effort and achieve better results.

These expectancy and value components jointly influence learners' academic outcomes, including their engagement, persistence, and performance (Gaspard, et al., 2015). For example, when learners expect to succeed and value the task highly, they are more likely to dedicate effort and achieve better results.

This theoretical framework sheds light on decision-making, perseverance, and motivation by emphasizing how individuals evaluate the costs and benefits associated with their actions. Motivation tends to increase when individuals feel confident in their ability to overcome challenges using their skills and strategies, especially when they perceive that the rewards of achieving their goals outweigh the effort or resources required. The Expectation-Value Theory has been widely applied across various disciplines, such as in physical education (e.g., Shang, Moss and Chen, 2023), language learning (e.g., Sun, et al., 2023), and mathematics (e.g., Fong, et al., 2023).

This theory suggests that students' intention to use ChatGPT is influenced by their perception of its value (e.g., utility, intrinsic interest, or alignment with goals) as well as the perceived cost (e.g., time, effort, or ethical concerns), with both factors playing critical roles in shaping motivation and decision-making. The integration of artificial intelligence applications in educational settings is seen as a method to improve learning outcomes, with students' perceptions playing a crucial role in their behavioral intentions and actual use (Chan and Zhou, 2023; Sankaran, et al., 2023). Li (2024) used the Situated Expectancy-Value Theory to study what influences pre-service EFL teachers' willingness to use technology, emphasizing the role of both expectancy beliefs and task value in how technology is adopted. This is relevant to the current study, which looks at how EFL learners' motivation to use generative AI in translation is shaped by how useful they think the tool is, how well it fits their learning goals, and what results they expect.

This study provides a chance to explore how Expectancy-Value Theory can help us better understand how students use AI tools in translation education, especially in the context of Vietnam. By examining students' views on AI's usefulness and its connection to their goals, we can gain insight into the psychological reasons behind their willingness to use generative AI in translation tasks.

# 2.4 Integrating ChatGPT and Expectation-Value Theory

By integrating insights from studies on generative AI and the Expectation-Value Theory, we can better understand how EFL students view the use of ChatGPT in translation classes. Learners who regard ChatGPT as an effective and engaging tool (high utility and intrinsic value) and who believe in their ability to use it successfully (high expectancy) are likely to be more motivated and perform better in translation tasks (Lo, et al., 2024; Slamet, 2024).

For instance, if students find that ChatGPT provides accurate translations and quick feedback, they are likely to see the tool as very useful. If they also enjoy using ChatGPT and find it interesting, the intrinsic value for the tool increases. When students believe that ChatGPT will help them succeed in their translation work, their confidence in doing well goes up, which boosts their motivation (Hmoud, et al., 2024).

This study builds on these ideas to look at how motivation affects students' views and behavior when using AI in translation. The framework guides both the analysis of numbers and personal experiences gathered in the research. It helps explore how students' views of AI's usefulness and their confidence in success shape their learning overall. The study aims to show how AI tools like ChatGPT can be added to language courses to improve both student engagement and learning results.

Specifically, this research investigates how EFL students feel about using ChatGPT in translation classes and how it affects their motivation and translation skills. To get a full picture, the study uses a mixed-methods approach that combines both surveys and interviews. Specifically, the objectives are as follows:

- To investigate students' perceptions of ChatGPT's value in translation tasks.
- To explore students' comparisons between their own translations and those generated by ChatGPT.

To address these objectives, two research questions were proposed:

RQ1: How do EFL students perceive the application of ChatGPT in translation tasks?

RQ2: How do students compare their manual translations with those generated by ChatGPT, and what factors influence their evaluation?

In this study, the term factors refers to the main variables that emerged from survey responses, reflective journals, and focus group discussions. These include how useful students found the tool, how fluent the translations were, how well cultural aspects were handled, as well as students' motivation and mental involvement. These factors align with the study's theoretical framework and research design. Using a mixed-methods approach helps combine different types of data, allowing the study to include both measurable results and students' personal reflections in the analysis.

#### 3. Method

#### 3.1 Research Design

This study employed a convergent parallel mixed methods approach, which combines both quantitative and qualitative data to offer a more comprehensive understanding of the research issue. The quantitative data were collected using pre- and post-survey questionnaires and end-course exam grades, while qualitative data were gathered through reflective journals and focus group interviews. In this approach, the data from both strands were analyzed separately and then compared to cross-check and deepen the understanding of the research problem (Creswell and Creswell, 2018).

The use of a mixed-methods design is justified in this study for several reasons. First, quantitative methods, through surveys and exam scores, allow for the measurement of students' attitudes and the impact of the intervention on their academic performance. Meanwhile, qualitative methods, such as reflective journals and focus groups, enable the researcher to explore the students' personal insights, experiences, and reflections regarding the use of generative AI tools in translation tasks. The combination of both types of data provides a more holistic view of how AI-assisted learning impacts students' motivation and performance.

On the first day of the Translation 2 course, the researcher administered pre-survey questionnaires to both Class 1 (control group) and Class 2 (experimental group). Afterward, Class 2 used generative AI (e.g., ChatGPT) to generate translations of a given text, comparing the AI-generated translations with their own translations. In contrast, Class 1 followed the traditional method, with the teacher providing corrections and feedback on students' translations. A post-survey questionnaire was administered to both groups after they completed their end-of-course exams. The effectiveness of the intervention was assessed by comparing the two groups' end-course grades, analyzing the pre- and post-surveys, and evaluating students' perspectives through their reflective journals and focus group interviews conducted at the end of the course.

#### 3.2 Participants

Sixty-two junior English majors from a private university in the Mekong Delta, Vietnam, participated in the study. The students were enrolled in two intact classes with the same curriculum, having studied "Translation 1" and now studying "Translation 2." The analysis of their pre-course translation performance, based on available grades, showed no significant differences between the two groups. Pre-course grades were analyzed using an independent samples t-test to assess any initial differences in performance between the two groups. The full statistical results can be found in the Findings section.

#### 3.3 Intervention: Students' Using ChatGPT for Their in-class Translation Practice

Classes 1 and 2 were assigned to translate a given Vietnamese text into English in groups, using only paper and pen to prevent them from using the Internet for assistance. Some technical terms were provided to aid their work. Afterward, Class 2 was allowed to use the free version of ChatGPT to translate the same text into English in class. The teacher instructed them to compare their translations with ChatGPT's version and take notes on anything they learned from the comparison.

Meanwhile, Class 1 followed a different approach: the teacher randomly selected translations from 2–3 groups, analyzed them, and corrected mistakes with the entire class. Both classes met with the teacher twice a week, and the same procedures were repeated until the end of the course.

#### 3.4 Research Instruments

This research employed pre-course and end-course grades (available on the university's website), survey questionnaires, reflective journals and focus group interviews. The questionnaire, adapted from Eccles and Wigfield (1995) and Zhu, et al. (2012), consisted of two sections: Section 1 surveyed students' demographic information, and Section 2 comprised an 11-item five-point Likert scale, with five items for expectancy beliefs measurement and six items for assessment of attainment, intrinsic, and utility values.

The next research tool was students' reflective journals or diaries. Reflective journals are particularly effective in capturing participants' time-related development, changes over time, and even potential causal relationships between variables (Dörnyei, 2007). This method encourages participants to document activities and reflections they consider meaningful, providing a rich source of qualitative data (Jacelon and Imperio, 2005). In this study, voluntary participants were asked to write self-reflective reports on their perceptions of the benefits and limitations of using ChatGPT to revise their groupwork translations. Guided questions such as "Do you think ChatGPT helps your translation version better for this week topic? How?" Or "Do you think ChatGPT makes your translation version worse? In what way it is worse?"

Participants were given the option to either write or audio-record their reflections on their computers and submit them through Google Forms, which were accessible exclusively to the researchers and the participants. They were encouraged to complete at least one reflective journal each week for a total of eight weeks. However, they had the flexibility to stop submitting reflections if they reported no new strategies to document.

The last research tool was focus-group interviews. Focus groups, involving participants who have experienced a specific situation, centralized the interaction among participants, rather than with the interviewer, with the interviewer serving as a moderator to facilitate discussion (Bryman, 2016). This dynamic can yield deep and insightful discussions (Dörnyei, 2007) and allow participants' views to emerge naturally (Cohen, Manion and Morrison, 2018). In this study, each group of Class 2 was invited to share their experiences of using ChatGPT for their translation tasks at their willingness. Eight out of ten groups agreed to participated in this phase of this data collection while the other two groups did not participate due to scheduling conflicts among their members.

The validity and reliability of the research instruments are supported through several measures outlined in the study. For the quantitative data, the internal validity of the questionnaire was ensured by (a) adapting items previously validated in earlier research (Eccles and Wigfield, 1995; Zhu, et al., 2012) and (b) conducting a piloting phase. Cronbach's Alpha values for the variables, all exceeding 0.7, indicated strong internal consistency (Table 2). Construct reliability was further demonstrated by consistent Cronbach's Alpha values across both the presurvey and post-survey phases (Table 1). For qualitative data, thematic analysis conducted by the researcher and a colleague followed the guidelines established by Braun and Clarke (2006), bolstering the validity of the findings. An inter-rater agreement rate of at least 75% (Mackey and Gass, 2022) was used as a benchmark, and in this study, an 80% agreement was achieved. Any inconsistencies in coding were managed through deliberation or removal from the analysis.

# 3.5 Data collection and Analysis

# 3.5.1 Piloting phase

To ensure the internal consistency of the survey items and assess respondents' comprehension, a pilot survey was conducted with 35 junior English major students from the same research site. The participants received an email in Vietnamese detailing the research purpose, estimated survey duration, and a request for their voluntary consent to participate. Their responses were automatically recorded in Google Sheets, accessible exclusively to the researchers.

# 3.5.2 Data collection procedures for the official research

The official survey questionnaires were distributed to 62 participants via email, with email addresses obtained from the university's publicly accessible staff directory. To ensure participants' willingness to participate, an email was sent explaining the research purpose and including a link to the questionnaire along with a participant consent form. Participants had a right to discontinue participation at any time without any repercussions. Presurvey data were collected from August 6 to 15, 2024, while post-survey data were gathered from October 15 to 22, 2024. The responses were automatically recorded in Google Sheets, accessible only to the researchers.

# 3.6 Reliability and Validity

The validity and reliability of the research instruments were ensured through several measures. For the quantitative data, the internal validity of the questionnaire was ensured by adapting items from previously validated instruments (Eccles and Wigfield, 1995; Zhu, et al., 2012) and conducting a piloting phase. Cronbach's Alpha values, indicating strong internal consistency, were above 0.7 for all variables.

Table 1: Construct reliability of the piloting phase

Variables	Cronbach's Alpha	N of items
Expectancy beliefs	.743	5
Attainment (Importance) value	.798	2
Intrinsic value	.707	2
Utility value	.725	2
Academic outcomes	.763	4

Table 2: Construct reliability of the actual phase

Variables	Cronbach's Alpha	N of Items
Expectancy beliefs	.716	5
Attainment (Importance) value	.815	2
Intrinsic value	.827	2
Utility value	.761	2
Academic outcomes	.908	4

# 4. Findings

# 4.1 RQ1: How do EFL students perceive the application of ChatGPT in translation tasks?

#### 4.1.1 Quantitative Results

Pre- and post-tests

The mean of two groups' end-course grades was calculated to compare the differences in their translation performance. Results from the pre-course analysis showed no significant difference between the two groups, as detailed below in Table 3 and Table 4.

Table 3: Mean scores of the two groups before the treatment

Descriptive Statistics								
Class	N	Mean	Std. Deviation	Std. Error Mean				
Class1	31	7.68	.86	.15				
Class2	31	7.38	.71	.12				

Table 4: Independent Samples t-Test Comparing Pre-course grades between groups

t	df	р	Mean diff.	SE Diff.	95% CI (Lower, Upper)
1.51	62	.136	.30	.20	-0.10, 0.70

**Note.** SE = standard error; CI = confidence interval

The next step was to compare the end-course exam grades to determine any differences in translation performance after the intervention. The post-course analysis showed that students in Class 2, who used ChatGPT as a tool for translation tasks, performed slightly better than those in Class 1, who followed the conventional approach. Table 5 below presents the comparison of the two groups' end-course exam grades.

Table 5: Comparison of the mean of the two groups' end-course exam grades

Group Statistics						
	Class	N	Mean	Std. Deviation	Std. Error Mean	
Post-tests	Class1	31	8.04	.46	.08	
	Class2	31	8.31	.58	.10	

The mean end-course grades of Class 1 (control group) was 8.035 (SD = 0.4550), while Class 2 (experimental group) achieved a higher mean score of 8.310 (SD = 0.5776). This difference suggests that students in Class 2, who were allowed to use ChatGPT, performed slightly better in their translation tasks compared to those in Class 1B, who followed a conventional approach without Al assistance. However, an Independent Samples T-Test was performed to assure that this difference was statistically significant or not (Table 6).

Table 6: Independent Samples t-Test Comparing Post-tests Scores Between Groups

t	df	р	Mean diff.	SE Diff.	95% CI (Lower, Upper)
-2.08	62	.042	-0.27	0.13	-0.54, -0.01

**Note.** SE = standard error; CI = confidence interval. Levene's test was not significant (p = .080), so equal variances were assumed. Table 6 showed that the groups' scores were statistically significant. The t-test showed a statistically significant difference in the average scores of the two classes' post-tests (t = -2.076, df = 62, p = 0.042).

The mean difference of -0.27 indicates that Class 2 scored about 0.27 points higher on average than Class 1. The 95% confidence interval for this difference ranged from -0.54 to -0.01, suggesting that while the difference is small, it is statistically meaningful.

Additionally, the effectiveness of the treatment was also calculated by the experimental groups' end-course performance. Hence, the pre-test and post-test of the treatment group were compared (Table 7):

Table 7: Descriptive Statistics for Pre- and Post-Test Scores in the Experimental Group

Group Statistics						
	Tests	N	Mean	Std. Deviation	Std. Error Mean	
Pre- and post-	Pre-test	31	7.38	.72	.13	
grades	Post-test	31	8.31	.58	.10	

To evaluate the impact of using ChatGPT, a paired-samples t-test was conducted to compare participants' scores before and after the intervention. The results are summarized in Table 8.

Table 8: Paired-Samples t-Test Comparing Pre- and Post-Test Scores

t	df	p	Mean diff.	SE Diff.	95% CI (Lower, Upper)
-5.59	62	< .001	-0.93	0.17	-1.26, -0.59

The comparison of pre- and post-test results reveals a significant improvement in the experimental group's performance after using ChatGPT. Table 8 presents the results of the independent samples t-test, which was conducted to assess the statistical significance of the difference between the two sets of scores.

The Levene's test confirmed that there was no significant difference in the variance between the pre- and post-test scores (p = 0.479), suggesting that the assumption of equal variances was not violated. The t-test results indicated a highly significant difference between the two sets of scores (p < 0.01). The mean post-test score was 8.310 (SD = 0.5776), which was higher than the pre-test score of 7.384 (SD = 0.7179), with a mean difference of -0.9258. This suggests that the experimental group's performance significantly improved after ChatGPT was incorporated into the translation process.

#### Pre- and Post-Surveys

To examine students' changes in their perceptions of using ChatGPT before and after the treatment, pre- and post-surveys were performed (Table 9). A paired t-test was performed to examine changes in students' perceptions of ChatGPT's effect on vocabulary and grammar before and after its use. This method was chosen because it accounts for the dependent nature of the data, as the same participants provided responses at two different time points (Field, 2018). The paired t-test is suitable for evaluating within-subject variations over time while controlling for individual differences (Pallant, 2020). Furthermore, it remains a valid statistical approach when the sample size is sufficiently large (typically n > 30-50), even if the data does not follow a normal distribution (Field, 2018).

Table 9: Paired-Samples t-Test Results for Pre- and Post-Surveys

Pair	t	df	р	Mean Diff.	SE	95% CI (Lower, Upper)
EB	-1.51	60	.136	-0.16	0.10	-0.37, 0.05
AV	-0.87	60	.387	-0.12	0.14	-0.41, 0.16
IV	-2.19	60	.033	-0.25	0.11	-0.47, -0.02
UV	-2.04	60	.046	-0.27	0.13	-0.54, -0.01
AA	-5.14	60	< .001	-0.67	0.13	-0.93, -0.41

The paired-samples t-test results provide a nuanced understanding of the changes in participants' scores between pre- and post-tests across different measures. While some variables demonstrated significant improvements, others remained unchanged.

Table 10 provides the descriptive statistics for participants' pre- and post-survey scores across five measured variables, showing the average values and variability at both time points.

Table 10: Descriptive Statistics for Pre- and Post-Survey Scores

Variable	Time point	n	Mean	SD	SE
EB	Pre	62	3.70	0.55	0.07
	Post	62	3.85	0.50	0.06
AV	Pre	62	3.91	0.84	0.11
	Post	62	4.03	0.69	0.09
IV	Pre	62	3.75	0.80	0.10
	Post	62	4.00	0.68	0.09
UV	Pre	62	3.78	0.79	0.10
	Post	62	4.05	0.74	0.09
AA	Pre	62	3.08	0.62	0.08
	Post	62	3.75	0.78	0.10

For instance, IV scores improved significantly from pre- to post-test (t(60) = -2.19, p = 0.033), with a small effect size (Cohen's d = -0.28). Similarly, UV scores showed a statistically significant increase (t(60) = -2.04, p = 0.046) with a small effect size (Cohen's d = -0.26). These findings suggest that participants experienced modest but measurable progress in these areas.

Table 11 presents the effect sizes for paired-samples comparisons, detailing the magnitude and confidence intervals of changes in each variable from pre- to post-test.

**Table 11: Effect Sizes for Paired-Samples Comparisons** 

Pair	Cohen's d	95% CI (Lower, Upper)	Hedges' g	95% CI (Lower, Upper
EB	-0.19	-0.45, 0.06	-0.19	-0.44, 0.06
AV	-0.11	-0.36, 0.14	-0.11	-0.36, 0.14
IV	-0.28	-0.54, -0.02	-0.28	-0.53, -0.02
UV	-0.26	-0.52, -0.01	-0.26	-0.51, -0.01
AA	-0.66	-0.93, -0.38	-0.65	-0.93, -0.38

**Note.** Cohen's d and Hedges' g represent standardized effect sizes. Negative values reflect post-test improvements. CI = confidence interval; EB = Expectancy Beliefs, AV = Attainment Value, IV = Intrinsic Value, UV = Utility Value, AA = Academic Achievement.

The most substantial improvement was observed in AA scores, which demonstrated a highly significant increase (t(60) = -5.14, p < 0.001) and a moderate effect size (Cohen's d = -0.66), demonstrating a meaningful change and the strongest effect of the intervention among all variables. However, EB scores (t(60) = -1.51, p = 0.136) and AV

scores (t(60) = -0.87, p = 0.387) showed no statistically significant differences, suggesting that the intervention did not substantially impact these measures.

These results highlight the targeted effectiveness of the intervention, particularly in improving AA scores, and to a lesser extent, IV and UV scores. The varying effect sizes indicate that while the intervention positively influenced some areas, its impact was not uniformly distributed across all measures.

#### 4.2 Qualitative Results

#### 4.2.1 Theme 1: Enhancing translation accuracy and vocabulary

Many participants perceived ChatGPT as a useful tool for the translation accuracy and helped them grow vocabulary through the use of contextually relevant word hints and insightful explanations for complexed or confusing epxressions. For instance, one student noted, "With ChatGPT, we can use the correct words for specific terms in the educational field, such as 'hoc ba' as 'school report' and "tuyển thẳng" as "direct recruitment'" (Student 7, Diary 1). Similarly, one significant advantage of ChatGPT was its handling of colloquial idioms, such as "nét bình dị đậm chất miền Tây" as "the rustic charm of the Mekong Delta" Group2\_Student 8). Another student valued the easy-to-understand explantions for technical terms so that they could make a right word choice for their context, noting, "It explains technical terms like the difference between 'submerge' and 'immerse'" (Student 18, Diary 1).

Students also appreciated the forte of ChatGPT in providing new and suitable synonyms, as well as broad lexical repertoire. One student took notes in a journal, "ChatGPT replaced 'emergency department' with a more specific term such as "Emergency, Intensive Care, and Toxicology", which helped improve clarity" (Student 4, Diary 2). Similarly, ChatGPT suggested a more accurate alternative like 'psychological trauma' instead of 'traumatic events' (Student 4, Diary 2). By offering contextual-based expressions, ChatGPT contributed to more natural translations (Student 9, Diary 1).

#### 4.2.2 Theme 2: Enhancing efficiency and fluency

Students commended ChatGPT's capacity to decrease the amount of time needed for revision while also enhancing the readability and fluidity of translations. For instance, one participant reflected, "ChatGPT helps me get translations faster, which helps me to understand the content better" (Student 15, Diary 4), while another added, "It processes and translates long sentences efficiently, and so I can save time for the task" (Group2\_Student 7).

Beyond providing rapid suggestions for translated text, ChatGPT's translations can help enrich sentence structure and reader-friendliness. A student shared, "ChatGPT transforms complex sentences into simpler and more coherent ones; this not only helps enhance readability but also improve sentence coherence" (Student 10, Diary 2). Additionally, it helped students avoid run-on sentences. "Its suggestions help reduce run-on sentences and improve sentence flow", shared Student 9 (Group2).

### 4.2.3 Theme 3: Enriching learning and collaboration

ChatGPT plays a role of a learning aid when expanding students' vocabulary and improving their sentence structures. Many participants found that ChatGPT supported their learning through topic-specific vocabulary. For instance, one student noted, "I learned precise terms like 'fields of study' instead of 'majors' in the phrase "Hanoi Pedagogical University 2 is recruiting about 2,000 students for 23 majors" and applied them to other contexts such as "Admission selection for fields such as Preschool Education, Physical Education" (Student 9, Diary 1). Another favored better terms that ChatGPT suggested for tourism topics, such as the expression of 'beautiful place' substituted by 'tourist spot' in "This tourist spot is beautiful thanks to nature and its altitude advantage. Here "điểm du lịch" literally means "tourist spot," a more precise term than just "beautiful place" (Student 12, Diary 2).

Furthermore, ChatGPT also facilitated group discussions and enhanced collaborative analysis. By comparing manual translations with ChatGPT's version and their manual work, students could identify errors and make improvements. One participant shared, "ChatGPT's translations help us analyze and understand different styles during group discussions. For example, one group's trasnslation of "Rừng già xếp tầng lớp dưới những khối mây có cảm giác như đất trời hòa làm một" was "The ancient forest is layered under clouds, which gives a feeling that earth and sky merge as one" which, I think, is better than ours "The ancient forest grows layers by layers under clouds which makes us have the feeling that earth and sky merge are one." (Group1\_Student 4).

# 4.2.4 Theme 4: Concerns about dependency

Despite its strengths, students were concerned about ChatGPT's influence on their possibility of being over-reliant on ChatGPT, which could thwart their critical thinking and empede their effort to unpack meanings from contexts. Some expressed this fear, "ChatGPT makes it easy to depend on its outputs, which can prevent me from developing my own skills" (Student 6, Diary 4). Another echoed this concern and added "Over-reliance on ChatGPT reduces the learning I gain from manual translations, sometimes I feel less confident about my [translated] version since I fear that my word choice or grammar is not academic enough" (Group2\_Student 8).

Beyond concerns about over-reliance on ChatGPT, students also faced other considerations, such as handling cultural and contextual nuances in the terms they used during their learning and translation techniques. For example, a participant pointed out, "ChatGPT struggles with cultural terms like "công đất" translating them awkwardly as "5 công of land" (Group1\_Student 5), or "miền Tây" translated as "the Western" (Group 1\_Student 6). Similarly, for idiomatic phrases, ChatGPT often missed the tone, requiring manual adjustments: "For idiomatic phrases, ChatGPT often misses the tone, requiring manual adjustments" (Student 13, Diary 4).

In addition to the themes identified from both the reflective journals and focus group interviews, additional themes emerged uniquely from each data collection method. Specifically:

# 4.2.5 Theme 5: Increasing confidence in translation (from reflective journals)

Although some students raised concerns about feeling less confidence in their manually translated version, most participants admitted a significant benefit of ChatGPT is its ability to enhance the quality of their translations for complex translation texts. By simplifying confusing and challenging phrases and terminologies, ChatGPT provided them with a more natural and contextual-bound terms, hence increased their confidence on the submitted tasks. For instance, a student was satisfied with the tool for its precise vocabulary for technical terms like 'psychological trauma,' 'toxicology,' and 'dermal injury,' which they had not encountered before (Student 4, Diary 2). Similarly, another student appreciated ChatGPT's ability to accurately translate specific terms such as 'post-traumatic stress disorder' as 'rối loạn căng thẳng sau chấn thương (PTSD)' and 'mechanical ventilation' as 'thở máy,' which enhanced their understanding and precision in medical translation (Student 9, Diary 2). This capability of ChatGPT not only improves the accuracy of targeted language translations but also encourages students to approach unfamiliar topics with greater confidence.

In addition to providing accurate terminology, ChatGPT helps reduce the anxiety often associated with tackling complex sentence structures, with several students acknowledging that using the tool eased their stress during translation tasks. For instance, one student shared, "Using ChatGPT makes me feel less overwhelmed when I have to turn long Vietnamese sentences into proper English, especially when I'm unsure how to arrange the ideas" (Student 13, Diary 4). Another student appreciated how the tool quickly breaks down complicated sentences into clear, manageable parts, helping them save time and reduce frustration. This allowed them to concentrate more on making their translations accurate (Student 15, Diary 4). By easing these difficulties, ChatGPT creates a more encouraging learning environment and boosts students' confidence as they improve their translation skills.

#### 4.2.6 Theme 6: Limitations in translation quality (from reflective journals)

While ChatGPT offers several advantages, interestingly, some students also identified significant limitations in its translation quality, particularly regarding the naturalness and accuracy of its outputs, especially when the terms related to cultural aspects. For example, a student noted that "ChatGPT's version 'The ancient forest is layered under clouds, which gives a feeling that earth and sky merge as one' sounds better than ours, but still feels a bit stiff and misses the poetic feel of the original Vietnamese description" (Group1\_Student 4). Similarly, another student reflected that sometimes ChatGPT's suggestions for overly academic or high-level vocabulary may not align with the intended tone of the translated text. For example, one diary noted "ChatGPT's word choices like 'altitude' and 'dense canopy' sound impressive and fancy, but don't always fit the relaxed and inviting mood of a travel description. This mismatch can make the translation feel less reader-friendly" (Student 17, Diary 4). These limitations suggest that while ChatGPT is effective for technical accuracy, it may lack the adaptability and cultural sensitivity required for more nuanced and contextually appropiate text-type style translations.

In addition to sounding overly formal, students reported instances of mistranslations and contextual errors that required manual intervention. For instance, one student remarked that "ChatGPT sometimes uses weird or incorrect words for the context," which can confuse the intended meaning (Student 5, Diary 2). For instance,

the homestay owner's activities like "bắt sò huyết" (catching cockles) or "bơi xuồng dỡ lú" (checking fish traps by boat) reflect local cultural aspects that ChatGPT might not render the naturalness or accuracy of the terms without intervention. Such mistranslations risk losing the nuance and authenticity crucial for effective communication, especially in community-based tourism contexts. Another noted that "some expressions don't sound natural and require manual adjustments," such as the phrase "Cung đường bám theo lưng chừng núi" (translated as "The route follows the mountainside"), which may sound technically correct but can lack natural fluency and vividness, hence needing human rephrasing for a more engaging and clear description. These findings emphasize the importance of critical engagement with ChatGPT's responses and require students' cultural knowledge and ability to refine its translations to align with the tone and context of the targeted translated text. While acknowledging the undeniable value of ChatGPT's support in text translation, translators should consider the aforementioned limitations to maintain a balance between Al assistance and manual translation efforts.

# 4.2.7 Theme 7: Facilitation of collaborative learning (from focus group interviews)

Many students from the focus group interviews said that owing to the suggested version from ChatGPT, their groups worked together again, compared their manual version, and then sharpened their final version before offcial submission to the teacher. These times when they felt most exploited from the ChatGPT to enrich knowledge as well as boosted teamwork. One participant shared, "We analyzed ChatGPT's translated outputs together to see whether they are correct, approriate or not," (Group1\_St4). Similarly, several emphasized that by comparing ChatGPT's with their manual translations, they developed the ability to identify errors and engaged in critical discussion to have a better version of their final product (Group 2\_St1). This process not only enhances students' translation skills but also promotes teamwork, as group members actively contribute their own perspectives and insights to refine their collective understanding of the required translation tasks.

In addition to strengthening collaborative analysis, ChatGPT aided thematic vocabulary building, hence further enriched the vocabulary learning experience. Students frequently used the tool to find out and document new vocabulary relevant to specific topics, which they afterwards incorporated into exams and other assignments. One student commented, "We noted down useful words by themes and applied them in our tests as well," indicating that ChatGPT can facilitate targeted vocabulary acquisition (Group1\_St2). By assisiting students in categorizing and retaining new vocabulary items systematically, ChatGPT enhances their lexical repertoire and prepares them for future translation tasks. Together, using ChatGPT in groupwork for translation tasks creates an interactive and dynamic learning environment where students harness ChatGPT's capabilities to optimize not only their individual performance but also their collaborative outcomes.

Figure 1 below illustrates these themes more vividly:

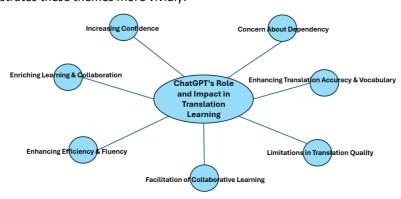


Figure 1: Students' perceptions of the application of ChatGPT in translation learning

# 4.3 RQ2: How do Students Compare Their Manual Translations With Those Generated by ChatGPT, and What Factors Influence Their Evaluation?

# 4.3.1 Theme 1: Comparative advantage of ChatGPT translations

Students consistently highlighted the positive side of ChatGPT since it can produce fluent, natural, and polished translations that often surpassed their manual ones in certain contexts. One student remarked, "ChatGPT is normally able to suggest academic words, and so its translation is with high level of naturalness and smoothness, which is, I think, closer to authentic papers" (Student 10, Diary 1). Similarly, another pointed out, "I like ChatGPT

because it helps with difficult phrases and makes the translations more concise and formal" (Student 8, Diary 4). In group discussions, participants emphasized the superior fluency of ChatGPT's outputs, "ChatGPT's translations feel smoother and more polished than our manual versions", one student acknowledged (Group2\_St10).

Another key advantage was the superiority of ChatGPT's provision of the variety of synonyms and diverse structures compared to students' manual outcomes. ChatGPT was able to provide alternative patterns and synonyms relevant to specific technical contexts, which helped facilitate the refinement of their self-translated versions. One participant shared, "ChatGPT gives various word options and diverse structures that we can consider the one that is most appropriate for the context" (Student 7, Diary 1). Another student appreciated ChatGPT's ability to offer appropriate technical terms such as 'psychological trauma' instead of 'traumatic events,' indicating that it could demonstrate a deeper, more specific contextual understanding (Student 4, Diary 2). This ChatGPT's capacity allowed students to explore and select the most suitable expressions to enhance the accuracy and contextual relevance of their translations.

#### 4.3.2 Theme 2: Preference for manual translations in specific contexts

Despite ChatGPT's advantages, students often preferred their manual translations in scenarios requiring cultural sensitivity or personal styles. Many expressed those manual efforts better captured nuanced meanings and cultural references. For example, one participant reported, "I sometimes see sentences [suggested by ChatGPT] that don't sound better than my own ones" (Student 18, Diary 2). Another added, "ChatGPT doesn't fully convey cultural nuances, especially idioms or specific topics related to local dialects" (Student 15, Group 4). This limitation was evident in examples like the mistranslation of '05 công đất' as '5 công of land,' which manual version can be much more accurate (Group2\_St9). Beyond limitations in obtaining cultural accuracy, several students found that they felt satified with their manual translations. One participant shared, "There are sentences where we feel our translations are better, so we don't use ChatGPT's version" (Student 4, Diary 3). Another remarked, "Given that AI models occasionally provide incorrect or misleading answers and tend to obscure my uncertainties, I am likely to favor my own translations", (Student 1\_Diary 4). These insights indicate that although ChatGPT are superior to students' manual translation in terms of technical and formal texts, manual translations remain crucial for capturing cultural and personal nuances.

#### 4.3.3 Theme 3: ChatGPT as a complementary learning tool

Alongside serving as a translation-assisted tool, the participants regarded them as a valuable resource for refining and improving their manual translations. Many participants emphasized its role in enhancing readability and sentence smoothness. For example, one student shared, "ChatGPT is helpful for fine-tuning translations, especially when I need more comprehensive expressions" (Student 10, Diary 2). Several are inclined to use it for improving and varying sentence structures, stating, "It helps improve the coherence of our translated versions by suggesting better structures and transition words" (Student 13, Diary 3). Additionally, ChatGPT was also useful for stylistic learning, such as providing variations in tone and phrasing. One participant said, "We asked ChatGPT for more formal expressions and described the context in which the sentence appears, then learned from its suggestions" (Group2\_St7).

Students characterized ChatGPT as a comparative learning tool that helped them identify and correct errors in their manual translations. One participant explained, "ChatGPT helps identify mistakes in our translations, such as typos, subject-verb agreement errors, or even the logical of thought. Thanks to this, we can fix them before submitting our paper to the teacher" (Group2\_St7). These reflections accentuate ChatGPT's potential as a complementary resource that supports students in fine-tuning their written products while promoting critical engagement with their work.

#### 4.3.4 Theme 4: Factors influencing translation evaluation

Some students acknowledged their shortcomings in advanced grammar and organizing ideas, and noted that using ChatGPT could help compensate for these weaknesses. ChatGPT has the cappability of producing grammatically and semantically coherent and cohesive sentences, which is conducive to students' learning of grammar and idea organization. One participant shared, "ChatGPT's alternative suggestions for our long sentences are helpful since they don't contain grammatical errors like run-on sentences and promote concise sentences that clearly convey multiple messages" (Group2\_St9). Additionally, ChatGPT's suggestions were often appreciated for their fluency and alignment with intended meanings. A participant stated, "ChatGPT translates more naturally and more closely in meaning to the targeted text than our versions, and even more so than Google Translate's" (Group1\_St2).

However, regarding unpacking meanings from some particular topics, especially Vienamese language when dialects exist among different regions, students indicated ChatGPT's limitations in translating this aspect. One student said, "I tried translating phrases like "bắt sò huyết" and "nhổ bồn bồn," but ChatGPT couldn't capture the local nuances; it translated them literally, which might confuse international readers' (Student 1, Group 8)." These evaluations indicate that while ChatGPT performs well in structured contexts, its effectiveness diminishes in informal or culturally nuanced scenarios, necessitating manual adjustments.

#### 4.3.5 Theme 5: Balancing AI and human effort

Appreciating ChatGPT's benefits and recognizing its limitations, many students expressed the importance of individual effort in text-based translation for personal development and the end-course exam, while emphasizing the value of integrating ChatGPT for consultation purposes. Many described ChatGPT as a valuable learning and refinement tool that enhanced the accuracy and clarity of their translations. For example, one student shared, "I usually ask ChatGPT to check my translated text; and it normally provides some key improvements that helps me refine my sentence" (Student 15, Diary 2). This reflects how students integrate ChatGPT into their learning process, using it as a complementary tool rather than a replacement one.

At the same time, students emphasized the importance of individuals' critical engagement and self-reliance in their translation process. Several students said that they normally do not use the ChatGPT's version without discussing in groups for the final version. One participant said, "We follow three steps: at first, we still translate manually; then we ask ChatGPT for help; and finally, we check ChatGPT's suggestions with teammates. By doing so, we can learn better" (Student 2, Group 7). Others prefered the opposite way, "whenever applicable, we put and paste the text to ChatGPT.com, then discuss its suggestion in groups and decide what to keep and what to remove from the sentence" (Student 1, Group 5). This mindset ensures that students remain actively engaged in the translation process while leveraging ChatGPT's capabilities to enhance their work.

#### 4.3.6 Theme 6: Balancing ChatGPT's strengths with student autonomy

Although the tempatation to fully using ChatGPT in their learning was huge, some students expressed concerns over its influence on their their learning autonomy and learning outcomes. One participant said, "If I just copy entire sentences or passages into ChatGPT, I'll gradually lose my ability to translate" (Student 3\_Group3). Others highlighted how important it is to balance manual work with Al-generated results. For example, a student noted, "We should only use ChatGPT as a reference tool to check clarity, make more concrete sentences, and not rely on it entirely" (Student 4\_Group1). Another emphasized the need to balance ChatGPT's use with their own critical thinking, "If I rely on ChatGPT for everything without making an effort to translate manually, my [translation] skills will erode, which is definitely detrimental to the final exam and my future as well" (Student2\_Group 7).

In summary, while ChatGPT is perceived as a valuable tool for improving translation accuracy, fluency, and vocabulary, students are also aware of its limitations, particularly in handling cultural nuances and idiomatic expressions. The tool can significantly enhance translation efficiency, support collaborative learning, and boost student confidence. However, concerns about dependency and the need for manual evaluation remain prominent. Ultimately, the findings suggest that ChatGPT is most beneficial when used as a supplementary tool that supports, rather than replaces, the manual translation process, and is not a one-size-fits-all solution. This balanced approach allows students to fine-tune their translations while retaining critical thinking skills, ensuring that they develop both their linguistic and analytical abilities.

Figure 2 below provides a clear illustration of these themes.

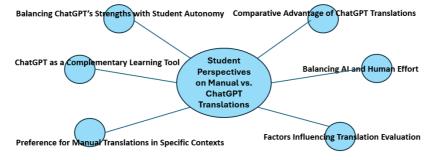


Figure 2: Student Perspectives on Manual vs. ChatGPT Translations

#### 5. Discussion

While the findings confirm the positive effect of integrating ChatGPT into translation tasks on students' translation performance, the overall impact of these improvements is modest. This raises thought-provoking questions regarding the practical ramifications of AI tools in language learning. For example, the small mean difference between pre- and post-test scores (a 0.27-point increase in the experimental group compared to the control one) suggests that factors such as task complexity, students' baseline competencies, and the degree of human evaluation or judgement might moderate the Generative AI tool's effectiveness. This aligns with the work of Sahari, AI-Kadi and AIi (2023), who found that while AI tools like ChatGPT can aid in mechanical translation tasks, their impact on overall learning outcomes can vary significantly across individuals and context-bound settings.

A critical discussion of the quantitative results reveals that although ChatGPT improves fluency and accuracy, its impact may be constrained by its inability to fully capture cultural nuances and informal language styles. This limitation is significant because, as noted by Sahari, Al-Kadi and Ali (2023), cultural sensitivity is crucial for high-quality translation. Therefore, while ChatGPT can act as an efficient support tool, educators must ensure that it is complemented by traditional teaching methods that develop students' critical thinking and contextual analysis skills.

Furthermore, the qualitative findings emphasize the roles of ChatGPT in fostering teamwork's learning collaboration and enriching students' understanding of translated texts. Students reported that comparing their manual translations with those generated by ChatGPT stimulated critical thinking and group discussions, which lends support to the findings by Ghassemiazghandi (2024) on the collaborative potential of AI in educational settings. This highlights that while ChatGPT can promote individual learning autonomy (Xiao and Zhi, 2023), it can also amplify benefits in group activities. The sharing and reflections from focus groups and reflective journals suggest that the foremost benefit of ChatGPT resides in its capacity to boost collaborative learning, especially in translation tasks where the translated outcomes require not only grammatical and semantic accuracy but also culturally embedded meaning. When students engage in discussions comparing AI outputs with their own work, they not only identify mechanical errors and issues related to meaning but also support each other in meaning negotiation.

While Van Horn's (2024) study emphasized ChatGPT's role in improving general language skills among Korean university students, this study provides specific insights into how ChatGPT supports task-specific learning, particularly translation. While Van Horn reported students' positive attitudes towards the potential of Al tools like ChatGPT, our findings delve deeper into students' reflective processes, supported with focus group interviews, showing how they integrate ChatGPT's suggestions with their manual efforts for better learning outcomes.

Furthermore, this study is consistent with a study by Sahari, Al-Kadi and Ali (2023) that highlighted ChatGPT's advantages in mechanical translation processes while admitting its limitations in processing cultural aspects and idiomatic expressions. However, our study advances this discourse by illustrating how students leveraged the tool as a supporting learning aid, enriching their understanding through group discussions, a dimension not previously explored in the previous study.

Unlike earlier studies that primarily used observational or survey methods (e.g., Sahari, Al-Kadi and Ali, 2023; Salloum, et al., 2024), this research employed a convergent parallel mixed-methods design, combining pre- and post-surveys, reflective journals, and focus group interviews. This methodological triangulation provided deeper insights into how students integrate ChatGPT into their translation practices and the resulting impact on their skills and perceptions. For instance, the significant improvement in students' post-test scores, alongside qualitative feedback, indicates that ChatGPT can enhance learning outcomes when used as a complementary tool rather than a standalone solution.

Additionally, the use of the Expectation-Value Theory (Wigfield & Eccles, 2000) to interpret students' perceptions offers a novel perspective. The findings reveal that students' motivation to use ChatGPT was influenced by its perceived utility, intrinsic value, and their expectancy for success. For instance, participants valued ChatGPT for its ability to simplify complex terms and reduce anxiety around translation tasks, which aligns with the theory's emphasis on task value and expectancy beliefs as motivators. This theoretical framing distinguishes the study from prior research, providing a more structured understanding of students' attitudes and behaviors.

#### 5.1 Contribution of the Current Study to the Existing Body of Literature

This study contributes to the growing body of research on the role of AI tools in language learning, particularly in translation education. By investigating EFL students' perspectives on using ChatGPT, the study expands upon existing findings and provides important insights into how AI tools can enhance learning outcomes and student engagement in translation tasks.

In line with prior studies (Xiao and Zhi, 2024; Van Horn, 2024), which emphasized the potential of AI in improving language accuracy and fostering autonomous learning, this research adds depth by exploring the intersection of AI use with self-regulated learning strategies. It builds on the work of Sahari, Al-Kadi and Ali (2023), who highlighted the benefits of AI tools for improving the mechanical aspects of translation (e.g., syntax and fluency). This study further confirms these findings while also emphasizing AI's limitations in addressing cultural nuances and idiomatic expressions, highlighting the need for human oversight in complex translation tasks.

Moreover, this study provides valuable insights for e-learning practices by showcasing how ChatGPT can be integrated into translation courses to enhance collaborative learning and critical thinking. Students found that using ChatGPT encouraged them to engage critically with the content, which resonates with generative Al's role in scaffolding learning (Karataş, et al., 2024). This collaborative element is crucial for the design of effective elearning environments, where Al serves as a supporting tool rather than a replacement for human interaction.

In practical terms, this research offers useful recommendations for educators considering AI tools in their curriculum. ChatGPT, as this study shows, can significantly reduce time spent on routine translation tasks, allowing students to focus more on content comprehension and higher-order translation skills. However, educators must ensure that AI is used in a way that fosters student autonomy and encourages critical engagement, thereby preventing over-reliance on AI outputs. This insight can inform curriculum design and assessment strategies in e-learning environments, ensuring that AI tools complement traditional learning methods rather than overshadow them.

#### 6. Conclusion

This study investigated English majors' perspectives on using text-based generative AI (specifically ChatGPT) in translation classes, employing the Expectation-Value Theory (Wigfield and Eccles, 2000) as a framework for understanding its impact on students' motivation, learning outcomes, and translation practices. The findings revealed that students valued ChatGPT as a valuable tool for providing timely and quality feedback on translation accuracy, efficiency, and vocabulary enrichment. In addition, ChatGPT plays a significant role in boosting students' motivation and confidence, and supporting their self-regulation and autonomous learning.

Despite these advantages, the study also identified critical limitations of ChatGPT, particularly in handling cultural nuances and idiomatic expressions. In these cases, students revised the suggested translated texts based on their teamwork discussions to accurately convey the meaning of both the source and target languages. A typical example of this is the translation of "05 công đất" or "miền Tây." These limitations highlight the importance of human review and critical engagement with Al-generated content. Therefore, while ChatGPT can mostly enhance translation fluency and accuracy, it should be considered and integrated as a complementary tool rather than a replacement for human expertise.

# 6.1 Strengths and Weaknesses of the Study

A major advantage of this study was its mixed-methods design, which combined a survey questuionnaire, preand post-tests from end-course exams, reflective journals, and focus group interviews to provide a comprehensive understanding of students' perceptions and experiences in utilizing ChatGPT during their translation classes. The pre- and post-tests objectively measured students' translation performance, while reflective journals and focus group interviews offered valuable insights into their learning processes. Additionally, the use of the Expectancy-Value Theory allowed for a structured analysis of students' motivation, contributing to the theoretical understanding of Generative-AI tools integration in higher education settings, particularly in translation skills practice.

However, certain limitations should be acknowledged. The study focused exclusively on English majors in Vietnam, which may limit the generalizability of the findings to other linguistic and cultural contexts. Additionally, the study's short-term intervention, lasting eight weeks and relying on pre- and post-tests, did not capture the long-term effects of Al-assisted translation learning.

#### 6.2 Future Directions for Research

Future research should focus on several key areas to expand our understanding of AI tools in translation education. First, due to the current study's limitation in short-term intervention, longitudinal studies could investigate the long-term effects of AI on translation skills and student motivation. Second, cross-cultural comparisons could broaden the scope of research by including diverse linguistic and cultural contexts. For example, research that includes students from different backgrounds would provide a more comprehensive understanding of AI's effectiveness in various educational settings. Lastly, this study highlights the important role of collaborative learning among team members but does not explore the potential of peer feedback from other groups. Therefore, future research could further investigate this issue to better understand social learning dynamics in translation education.

**Al Statement:** The authors declare that Artificial Intelligence tools were not used in this study for the paper's conception, revision, or the creation of figures and tables.

**Ethics Statement:** Prior to the data collection, ethical approvals have been obtained from all participants.

#### References

- Alawida, M., Mejri, S., Mehmood, A., Chikhaoui, B. and Isaac Abiodun, O., 2023. A comprehensive study of ChatGPT: advancements, limitations, and ethical considerations in natural language processing and cybersecurity. *Information*, [e-journal] 14(8), pp.1 23. Available at: <a href="https://www.mdpi.com/2078-2489/14/8/462">https://www.mdpi.com/2078-2489/14/8/462</a> [Accessed 03 June 2024].
- Barrot, J., 2024. ChatGPT as a language learning tool: An emerging technology report. *Technology, Knowledge and Learning*, [e-journal] 29(2), pp.1151 1156. https://doi.org/10.1007/s10758-023-09711-4.
- Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, [e-journal] 3(2), pp.77 101. https://doi.org/10.1191/1478088706qp063oa.
- Bryman, A., 2016. Social research methods. 5th ed. Oxford: Oxford University Press.
- Canning, E.A. and Harackiewicz, J.M., 2015. Teach it, don't preach it: The differential effects of directly-communicated and self-generated utility-value information. *Motivation science*, 1(1), 47. <a href="https://doi.org/10.1037/mot0000015">https://doi.org/10.1037/mot0000015</a>
- Chan, C.K.Y. and Zhou, W., 2023. An expectancy value theory (EVT) based instrument for measuring student perceptions of generative Al. *Smart Learning Environments*, [e-journal] 10(1), pp.1 22. <a href="https://doi.org/10.1186/s40561-023-00284-4">https://doi.org/10.1186/s40561-023-00284-4</a>.
- Cohen, L., Manion, L. and Morrison, K., 2018. Research methods in education. 8th ed. New York: Routledge.
- Creswell, J.W. and Creswell, J.D., 2018. Research design: Qualitative, quantitative, and mixed methods approaches. 5th ed. London: Sage.
- Dörnyei, Z., 2007. Research methods in applied linguistics. Oxford: Oxford University Press.
- Eccles, J.S. and Wigfield, A., 1995. In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, [e-journal] 21(3), pp.215 225. <a href="https://doi.org/10.1177/0146167295213003">https://doi.org/10.1177/0146167295213003</a>.
- Field, A., 2018. Discovering statistics using IBM SPSS statistics. 5th ed. California: SAGE Publications.
- Flake, J.K., Barron, K.E., Hulleman, C., McCoach, B.D. and Welsh, M.E., 2015. Measuring cost: The forgotten component of expectancy-value theory. *Contemporary educational psychology*, [e-journal] 41, pp.232-244. <a href="https://doi.org/10.1016/j.cedpsych.2015.03.002">https://doi.org/10.1016/j.cedpsych.2015.03.002</a>
- Fong, C.J., Lee, J., Krou, M.R., Hoff, M.A., Johnston-Ashton, K., Gonzales, C. and Beretvas, S.N., 2023. Meta-analyzing the factor structure of the learning and study strategies inventory. *Journal of Experimental Education*, [e-journal] 91(2), pp.380 400. https://doi.org/10.1080/00220973.2021.2021842.
- Gaspard, H., Dicke, A.L., Flunger, B., Brisson, B.M., Häfner, I., Nagengast, B. and Trautwein, U., 2015. Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental psychology*, [e-journal] 51(9), pp.1226 1240. <a href="https://doi.org/10.1037/dev0000028">https://doi.org/10.1037/dev0000028</a>
- Ghassemiazghandi, M., 2024. An evaluation of ChatGPT's translation accuracy using BLEU score. *Theory and Practice in Language Studies*, [e-journal] 14(4), pp.985 994. https://doi.org/10.17507/tpls.1404.07
- Harackiewicz, J.M., Smith, J.L. and Priniski, S.J., 2016. Interest matters: The importance of promoting interest in education. *Policy insights from the behavioral and brain sciences*, *3*(2), pp.220 227. https://doi.org/10.1177/2372732216655542
- Hmoud, M., Swaity, H., Hamad, N., Karram, O. and Daher, W., 2024. Higher education students' task motivation in the generative artificial intelligence context: The case of chatgpt. *Information*, [e-journal] 15(1), pp.1 18. Available at: <a href="https://pdfs.semanticscholar.org/f010/9696f537e6d59070bdeb9594ceac38925b70.pdf">https://pdfs.semanticscholar.org/f010/9696f537e6d59070bdeb9594ceac38925b70.pdf</a> [Accessed 03 June 2024]
- Jacelon, C.S. and Imperio, K., 2005. Participant diaries as a source of data in research with older adults. *Qualitative health research*, [e-journal] 15(7), pp.991 997. https://doi.org/10.1177/1049732305278603.
- Karataş, F., Yaşar, F. and Gunyel, F.O., 2024. Incorporating Al in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, [e-journal] 29(15), pp.19343 19366. https://doi.org/10.1007/s10639-024-12574-6.

- Li, S., 2024. Exploring pre-service foreign language teachers' intention to adopt technology: An investigation from the perspective of situated expectancy-value theory. *Journal of China Computer-Assisted Language Learning* [e-journal]. <a href="https://doi.org/10.1515/jccall-20240012">https://doi.org/10.1515/jccall-20240012</a>.
- Liu, G.L. and Ma, C., 2023. Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, [e-journal] 18(2), pp.125 138. <a href="https://doi.org/10.1080/17501229.2023.2240316">https://doi.org/10.1080/17501229.2023.2240316</a>.
- Lo, C.K., Yu, P.L.H., Xu, S., Ng, D.T.K. and Jong, M.S.Y., 2024. Exploring the application of ChatGPT in ESL/ EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, [e-journal] 11(1), pp.1 24. https://doi.org/10.1186/s40561-024-00342-5.
- Mackey, A. and Gass, S.M., 2022. Second language research: Methodology and design. 3rd ed. New York: Routledge.

  McIntosh, T.R., Susnjak, T., Arachchilage, N., Liu, T., Xu, D., Watters, P. and Halgamuge, M.N., 2025. Inadequacies of large language model benchmarks in the era of generative artificial intelligence. IEEE Transactions on Artificial Intelligence. <a href="https://doi.org/10.1109/TAI.2025.3569516">https://doi.org/10.1109/TAI.2025.3569516</a>
- Pallant, J., 2020. SPSS Survival Manual: A step by step guide to data analysis using IBM SPSS. 7th ed. New York: Routledge. Ruoqi, S., Yuan, X. and Gochuico, M.T., 2023. Translators and their use of ChatGPT. Academia Lasalliana Journal of Education and Humanities, [e-journal] 5(1), pp.49 59. https://doi.org/10.55902/PLRU9197.
- Sahari, Y., Al-Kadi, A.M.T. and Ali, J. K. M., 2023. A cross sectional study of ChatGPT in translation: Magnitude of use, attitudes, and uncertainties. *Journal of Psycholinguistic Research*, [e-journal] 52(6), pp.2937 2954. <a href="https://doi.org/10.1007/s10936-023-10031-y">https://doi.org/10.1007/s10936-023-10031-y</a>.
- Salloum, S.A., Aljanada, R.A., Alfaisal, A.M., Al Saidat, M.R. and Alfaisal, R., 2024. Exploring the acceptance of ChatGPT for translation: An extended TAM. In: A. Al-Marzouqi, A. A. Salloum, M. Al-Saidat, A. Aburayya, and B. Gupta, eds. 2024. *Artificial Intelligence in Education: The Power and Dangers of ChatGPT in the Classroom*. Cham: Springer Nature Switzerland, pp.527 - 542. <a href="https://doi.org/10.1007/978-3-031-52280-2">https://doi.org/10.1007/978-3-031-52280-2</a>.
- Sankaran, P., Deshbhag, R., Durbha, K., Gururajan, R. and Zhou, X., 2023. Student perceptions of ChatGPT through an expectancy value theory. In: 2023 IEEE International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT). Venice, Italy, 26-29 October 2023. London: IEEE. <a href="https://doi.org/10.1109/WI-IAT59888.2023.00089">https://doi.org/10.1109/WI-IAT59888.2023.00089</a>.
- Shang, C., Moss, A.C. and Chen, A., 2023. The expectancy-value theory: A meta-analysis of its application in physical education. *Journal of Sport and Health Science*, [e-journal] 12(1), pp.52 64. <a href="https://doi.org/10.1016/j.jshs.2022.01.003">https://doi.org/10.1016/j.jshs.2022.01.003</a>.
- Slamet, J., 2024. Potential of ChatGPT as a digital language learning assistant: EFL teachers' and students' perceptions. Discover Artificial Intelligence, [e-journal] 4(1), pp.1 - 15. https://doi.org/10.1007/s44163-024-00143-2.
- Solak, E., 2024. Revolutionizing language learning: How ChatGPT and AI are changing the way we learn languages. International Journal of Technology in Education, [e-journal] 7(2), pp.353 - 372. Available at: <a href="https://eric.ed.gov/?id=EJ1426851">https://eric.ed.gov/?id=EJ1426851</a> [Accessed 17 June 2024]
- Sun, W., Hong, J.C., Dong, Y., Huang, Y. and Fu, Q., 2023. Self-directed learning predicts online learning engagement in higher education mediated by perceived value of knowing learning goals. *Asia-Pacific Education Researcher*, [e-journal] 32(3), pp.307 316. <a href="https://doi.org/10.1007/s40299-022-00653-6">https://doi.org/10.1007/s40299-022-00653-6</a>.
- Trautwein, U., Lüdtke, O., Marsh, H.W. and Nagy, G., 2009. Within-school social comparison: How students perceive the standing of their class predicts academic self-concept. *Journal of educational psychology*, [e-journal] 101(4), pp.853 866. https://doi.org/10.1037/a0016306
- Van Horn, K.R., 2024. ChatGPT in English language learning: Exploring perceptions and promoting autonomy in a university EFL context. *TESL-EJ*, [e-journal] 28(1), pp.1 26. Available at: <a href="https://eric.ed.gov/?id=EJ1433988">https://eric.ed.gov/?id=EJ1433988</a>> [Accessed 07 June 2024]
- Wigfield, A. and Eccles, J.S., 2000. Expectancy value theory of achievement motivation. *Contemporary educational psychology*, [e-journal] 25(1), pp.68 81. <a href="https://doi.org/10.1006/ceps.1999.1015">https://doi.org/10.1006/ceps.1999.1015</a>.
- Wigfield, A., Tonks, S.M. and Klauda, S.L. (2016). Expectancy-value theory. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school.* 2nd ed., pp.55 74). Routledge.
- Xiao, Y. and Zhi, Y., 2023. An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, [e-journal] 8(3), pp.1 12. <a href="https://doi.org/10.3390/languages8030212">https://doi.org/10.3390/languages8030212</a>.
- Zhu, X., Sun, H., Chen, A. and Ennis, C., 2012. Measurement invariance of expectancy-value questionnaire in physical education. *Measurement in Physical Education and Exercise Science*, [e-journal] 16(1), pp.41 54. Available at: <a href="https://www.tandfonline.com/doi/abs/10.1080/1091367x.2012.639629">https://www.tandfonline.com/doi/abs/10.1080/1091367x.2012.639629</a> [Accessed 27 June 2024]