Journal of Languages and Translation (*JLT*)
Vol. 12, Issue 1, pp. 67-95 | January 2025
Print ISSN: 2735-4520 | E-ISSN: 2735-4539
https://jltmin.journals.ekb.eg/

Evaluating English-Arabic translation: Human translators vs. Google Translate and ChatGPT

Saudi Sadiq

Orcid.org/0000-0003-4491-4130

Department of English, Faculty of Al-Alsun (Languages), Minia University, Egypt
Email: saudi.sadiq@mu.edu.eg

Abstract

Machine Translation (MT) has seen rapid advancements due to progress in neural sciences and Large Language Models (LLMs). While human translation remains highly competitive across many language pairs, especially those that are linguistically and culturally close like English and Spanish (Moslem et al., 2023), the situation between English and Arabic requires further exploration. This study investigates the performance of Neural Machine Translation (NMT) represented by Google Translate, LLMs represented by ChatGPT, and human translation in translating English texts into Arabic across four genres: general, literary, scientific, and media. Human evaluations were used to measure translation quality based on accuracy, fluency, style, cultural suitability, and terminology. The results indicate that human translation is the most accurate, especially in capturing cultural and contextual nuances. ChatGPT, when used with detailed prompts, often outperforms both Google Translate and ChatGPT with simple prompts, particularly in literary and media genres. On the contrary, Google Translate performs the worst overall, especially with scientific and general texts, due to issues like word confusion and cultural inaccuracies. These results offer practical insights for translation educators, professionals, and students on effectively integrating MT tools while appreciating the irreplaceable value of human translators.

Keywords: ChatGPT, Google Translate, human translation, Machine Translation, translation quality evaluation

1. Introduction

MT refers to using computer algorithms to convert text from one language to another (Hutchins and Somers, 1992). These algorithms have evolved over time. Rule-Based Machine Translation (RBMT) relies on linguistic rules and dictionaries but requires extensive manual setup and lacks flexibility (Okpor, 2014). Statistical Machine Translation (SMT) uses probabilistic models built from large corpora to align words and phrases; however, its effectiveness depends on data quality, and it struggles with complex phrases (Hearne and Way, 2011). Neural Machine Translation employs deep learning and neural networks to analyze the meaning of the source text, focusing on important parts using attention mechanisms to create accurate and

context-aware translations (Bahdanau et al., 2014; Zong, 2018). While NMT generally outperforms SMT, both approaches can be combined in Hybrid Machine Translation (HMT) for improved results (Bahdanau et al., 2014). Example-Based Machine Translation (EBMT) uses previous translations to enhance accuracy but may struggle with unfamiliar texts (Hutchins, 2005).

Recent advancements in LLMs, such as ChatGPT, Gemini, Claude, Copilot, and Mistral AI, have significantly improved MT capabilities, even though these models were not specifically designed for translation. With the development of NMT and LLMs, MT has become widely used for standard or technical translations, particularly in contexts with minimal cultural nuance. Platforms like Amazon and social media sites such as Facebook and X rely on MT for localization and user interactions across languages. However, for texts rich in cultural or context-specific content, human translators may still be preferred, such as for legal documents, literary works, or marketing materials tailored for local markets. Even in these cases, translators often start with machine-generated translations and refine them for accuracy and cultural suitability.

Google Translate, launched in 2006 as an SMT tool, transitioned to the Google Neural Machine Translation (GNMT) engine in 2016. This upgrade improved its ability to understand context, extract relevant information, and infer meanings, significantly enhancing translation accuracy through deep learning techniques. Consequently, it began producing higher-quality translations. Human evaluations showed that the new system reduced errors by nearly 60% (Ducar and Schocket, 2018). The quality of translations varies considerably between language pairs. Languages with extensive data, such as English, Spanish, and French, benefit from more accurate and context-aware translations. In contrast, languages or dialects with fewer available resources tend to have lower translation quality (Aiken, 2019; Khoong et al., 2019; Taira et al., 2021). This is likely why Google Translate allows users to provide feedback by rating translations and suggesting edits, which may contribute to improving accuracy and lead to more reliable results over time. By June 2024, Google Translate supported translations in over 243 languages and varieties, making it one of the most widely used translation tools globally (Caswell, 2024).

ChatGPT, introduced in 2022 and now available in its latest version (ChatGPT-4o), is based on OpenAI's LLM technology. This pre-trained AI model is designed for natural language interactions and utilizes advanced methods from Natural Language Processing (NLP), Supervised Learning, and Reinforcement Learning to generate text that closely resembles human writing (Roumeliotis and Tselikas, 2023). Pre-training on extensive text data equips LLMs (e.g., ChatGPT) to grasp language nuances and generate accurate responses, even in complex or ambiguous contexts (Radford et al., 2019). Although ChatGPT is not specifically designed for translation, it is often used for translation between languages and language varieties. By using carefully crafted prompts, such as providing context or specifying the tone and style, ChatGPT can produce more accurate and relevant translations. This approach helps ChatGPT apply its pre-trained knowledge effectively to various translation needs (Gánem, 2024; Nagi et al., 2024). Similar to Google Translate, ChatGPT's translation quality varies by language pair. It provides more accurate translations for languages with extensive data, like

English and Spanish, but may be less reliable for languages with fewer resources (Gánem, 2024).

Comparing Google Translate and ChatGPT allows users to understand their strengths and limitations, helping them choose the most appropriate tool for their translation needs. Also, comparing MT to professional human translations—which are often more accurate and coherent for complex texts—is crucial. It is hypothesized that this comparison highlights areas where MT performs well and where it still falls short.

2. Objectives and research questions

This study compares MT tools—Google Translate (NMT) and ChatGPT (LLMs)—with human translation across various text types: general, literary, scientific, and media. These categories include the most frequently translated content from English into Arabic. Based on this background, the study aims to address the following questions:

1. How do Google Translate and ChatGPT perform in comparison to human translation across different text types (general, literary, scientific, and media) when translating from English into Arabic?
2. What are the strengths and weaknesses of Google Translate and ChatGPT compared to human translation for general, literary, scientific, and media texts?
3. Does varying the prompts enhance translation quality in ChatGPT?

3. Literature Review

3.1. Comparative performance of LLMs, NMT, and human translation

Research on using LLMs as MT tools is limited due to their recent adoption for this purpose. However, the inconsistent results when comparing LLMs to NMT tools in MT are noteworthy. In some instances, LLMs have shown better performance than NMT tools. For example, Alkhawaja (2024) compared ChatGPT and Google Translate in translating 1,000 English sentences sourced from Tatoeba, a collaborative online database of sentences and translations, into Arabic and found that ChatGPT performed slightly better than Google Translate. Similarly, Aldawsari (2024) found that Bing AI Chat outperformed both Google Translate and Bing Translator when translating Arabic colloquial expressions. Conversely, Banimelhem and Amayreh (2023) evaluated ChatGPT's English-to-Arabic translation against fifteen commercial MT systems, including Google Translate, and concluded that ChatGPT's performance on 100 English sentences from a standardized dataset was inferior to the other systems.

The quality of LLMs translations can vary greatly across different tools when applied to the same dataset. Algobaei et al. (2024) evaluated the translation of gender-sensitive terms by ChatGPT and Gemini, finding that Gemini outperformed ChatGPT. Comparisons often indicate that human translation surpasses LLMs in quality due to greater cultural sensitivity and contextual understanding. For example, Alkhawaja (2024) found that while ChatGPT slightly outperformed Google Translate, human translators surpassed both. Similarly, Moneus and Saharib (2024) compared the translations of legal contracts from English to Arabic performed by human translators and various LLMs. They concluded that while LLMs offer

quicker and lower-cost solutions, they often lack the cultural and contextual accuracy essential for legal documents.

Despite significant improvements since its transition from SMT to NMT in 2016, Google Translate has faced various criticisms, particularly regarding its effectiveness. It struggles with Arabic lexical ambiguity, including issues with homonyms and polysemies, which negatively affects translation quality (Aldawsari, 2023). It also frequently makes errors, especially when translating relative clauses from English into Arabic (Nagi, 2023). The tool also translates idioms word-for-word and renders colloquial Arabic based on words that look similar in Standard Arabic, leading to inaccuracies (Al-Sabbagh, 2024). Google Translate was also found to lack adherence to key legal norms, failing to effectively handle passivization, modality, collocations, and doublets, thus reducing the legal accuracy and quality of its translations compared to human translation (Mohamed, 2023). Earlier, Daniele (2019) quantitatively evaluated Google Translate's error rate in translating medical texts from English to Italian, identifying an overall error rate of 15%.

The cultural aspect of MT presents significant challenges, particularly in preserving the nuances of literary and idiomatic content. Alowedi and Al-Ahdal (2023) observed that MT tools often fail to preserve the cultural and lyrical elements of Arabic poetry. Similarly, Al-Kaabi et al. (2024) evaluated ChatGPT and Google Translate's ability to convey cultural meanings in Naguib Mahfouz's *Zuqāq al-Midaqq* and found that, although ChatGPT performed better, it still did not reach the quality of human translation in preserving cultural nuances. El-Saadany (2024) examined ChatGPT's translation of Arabic proverbs into English, highlighting its difficulty in maintaining original cultural meanings. Likewise, Al-Khresheh and Almaaytah (2018) noted Google Translate's limitations with English proverbs, attributing inaccuracies to issues like word ambiguity and structural differences. Taleghani and Pazouki (2018) also found that various online MT tools produced poor translations of English idioms and phrasal verbs into Persian, often missing essential cultural context.

3.2. Translation quality evaluation

Translation quality evaluation has been a debated topic for a long time, with many different opinions and methods offered by scholars in the field. Many of these perspectives have relied on vague and impressionistic criteria, which can lead to inconsistent evaluations of translations. Al-Qinai (2000) offers a critical analysis of these earlier models, highlighting their reliance on subjective evaluations that often lack clarity and precision. He points out that such approaches fail to provide a comprehensive understanding of translation quality, as they do not adequately consider the complexities involved in the translator's decision-making process.

In response to these limitations, Al-Qinai (2000) draws upon the foundational work of notable scholars, including Newmark (1988), Hatim and Mason (1990), and House (1981), to develop a more robust framework for evaluating translation quality. His multi-dimensional model moves beyond the impressionistic evaluations that have characterized much of the previous discourse. Instead, it emphasizes a systematic evaluation process that incorporates specific criteria across three key dimensions: textual, linguistic, and pragmatic, including:

- *Textual typology* involves analyzing the function and type of the text.

- *Formal correspondence* facilitates the comparison of structural elements like sentence length and paragraph formatting.
- *Coherence* focuses on the logical progression of information, particularly in managing thematic structures during translation.
- *Cohesion* ensures clarity by effectively using conjunctions, pronouns, and repetition.
- *Lexical properties* evaluate the accuracy of word choices, including the translation of idioms, metaphors, and culturally specific expressions.
- *Syntactic equivalence* guarantees the preservation of sentence structure and word order without distorting meaning.
- *Pragmatic equivalence* evaluates whether the translation produces a similar effect on the target audience, taking into account cultural nuances and the intended purpose of specific expressions.

4. Methodology

To address the research questions, a questionnaire was created featuring four texts representing the text types under investigation: a general text (215 words) from Said's 1981 *Covering Islam*, a literary text (219 words) from Morpurgo's 1999 children novel *Kensuke's Kingdom*, a scientific text (220 words) from Suppakittpaisarn et al. (2019), and a media text (143 words) from BBC (Yousif, 2024a). Three highly educated native English speakers were asked to evaluate these texts as complete texts. They reached a consensus that the texts were well-structured, with coherent and comprehensible content. Several texts were replaced to ensure that the final selection consisted of concise texts conveying complete and clear ideas.

The questionnaire (Appendix 1) involved evaluating four translations of each text. The first was done by Google Translate, the second and third by the free version of ChatGPT 4o—the second using the simple prompt "Translate the following text from English into Arabic" and the third using detailed prompts (see Table 1 for the prompts). The fourth is a human translation, with the general text translated by Muhammad Enani[1] (Said, 2022), the literary text also by Muhammad Enani (Morpurgo, 2023), the scientific text by the author, and the media text by BBC Arabic (Yousif, 2024b).

Research shows that prompt quality directly affects output quality, with specific prompts improving results (Giray, 2023). Although prompt engineering remains relatively unexplored in translation studies, several effective strategies have emerged. Jiao et al. (2023) introduced the pivot strategy, which involves translating into a high-resource language before the target language, leading to significant improvements. Gao et al. (2023) found that incorporating task details, domain context, or part-of-speech tags in prompts enhances ChatGPT's translation performance. Delta (2024) expanded on this by incorporating persona and tone into the prompts. Gu (2023) demonstrated that linguistically informed prompts can increase translation accuracy by over 35%. Guided by Gao et al. (2023) and Delta (2024), the third translation of each text was completed using ChatGPT 4o with detailed prompts (henceforth ChatGPT-D)

[1] Muhammad Enani is a prominent Egyptian scholar and translator known for his contributions to translation studies, particularly in translating Arabic and English. He is recognized for his translations of Arabic literature and his academic work on translation theory, focusing on the challenges of translating cultural and literary texts.

provided in Table 1. These prompts included specifications for the task, persona, context, and tone, allowing for a comparison with the second translation, which was generated using a simpler, task-only prompt (henceforth ChatGPT-S).

Ten translation specialists, all native Arabic speakers and experts in English-Arabic translation, were approached to serve as human evaluators for the questionnaire. While automatic metrics of translation evaluation, such as Bilingual Evaluation Understudy (BLEU)[2], Character n-gram F-score (chrF)[3], and Translation Edit Rate or Translation Error Rate (TER)[4], provide quick evaluations, they often fall short in capturing the nuanced aspects of translation quality, including fluency, cultural context, and meaning preservation (Son and Kim, 2023). These metrics primarily measure text similarity at a superficial level, conflating fluency with content accuracy (Jiang and Zhang, 2024). In contrast, human evaluations take into account a broader range of factors, such as the intended audience, cultural sensitivity, translation norms, coherence, and practicality. As a result, human evaluation is considered the most reliable method for evaluating translation quality (Papineni et al., 2022). For this reason, the current study employed professional translators as human evaluators. To avoid bias, evaluators were kept unaware of which translations were generated by Google, ChatGPT, or human translators.

The questionnaire used to evaluate translation quality is based on the systematic criteria described in Al-Qinai's (2000) model, explained in Section 3.2 above. It evaluates translations in five key areas that match his framework: accuracy in expressing the meaning of the source text, clarity and ease of understanding, matching the original style and tone, cultural suitability for the target audience, and correct use of specialized terms. Evaluators answered the questionnaire using a 5-point Likert scale, from 'Completely 5' to 'Not at all 1' for each translation. They were also encouraged to give extra comments, allowing for more detailed feedback, which helped to enrich the analysis of the results.

Table 1. Prompts utilized in ChatGPT

Text	**ChatGPT-S**	**ChatGPT-D**
Text 1: General	Translate the following text into Arabic	As a translator specializing in English-to-Arabic general translation, translate the following extract from Edward Said's 1981 book *Covering Islam* into Modern Standard Arabic. Ensure the translation is accurate, culturally appropriate, and suitable for an Arabic-speaking audience in the Middle East. Maintain the tone and style of the original text while adjusting the language to be appropriate and engaging for a general audience. Pay attention to consistency in key terms, especially names and recurring concepts, and restructure the syntax where necessary to ensure clarity and a natural flow in Arabic. Avoid overly literal translations, focusing instead on creating a smooth and coherent text.

[2] BLEU is a metric for automatically evaluating MT output by measuring its similarity to a set of high-quality reference translations using n-gram precision. The BLEU score ranges from 0 to 1, where 0 indicates no overlap (low quality) and 1 indicates perfect overlap (high quality) between the machine-generated translation and the references (Papineni et al., 2022).

[3] chrF is an MT evaluation metric that measures the overlap of character sequences (n-grams) between a machine translation and a reference. It is less affected by word splitting and gives partial credit for partly correct words (Popović, 2015).

[4] TER is an automatic evaluation metric for MT that measures the number of edits needed to modify a machine translation output to match a reference translation (Snover et al., 2006).

		Use clear and simple language, refraining from technical or specialized terminology unless necessary for the meaning of the text.
Text 2: Literary	Translate the following text into Arabic	As a translator specializing in English-to-Arabic literary translation, translate the following extract from Morpurgo's 1999 children's novel *Kensuke's Kingdom* into Modern Standard Arabic. Ensure the translation is accurate, culturally appropriate, and suitable for a young audience in the Middle East. Maintain the tone and style of the original text while adjusting the language to be age-appropriate and easy to read. Pay attention to consistency in key terms, particularly names and recurring concepts, and restructure the syntax where necessary to ensure clarity and a natural flow in Arabic, avoiding awkward or overly literal translations. Focus on creating a smooth, engaging, and accessible text for children, using simple language without resorting to technical or specialized terminology unless required by the narrative.
Text 3: Scientific	Translate the following text into Arabic	As a translator specializing in English-to-Arabic academic translation, translate the following abstract from an academic paper titled "Preferences for Green Infrastructure and Green Stormwater Infrastructure in Urban Landscapes: Differences between Designers and Laypeople" into Modern Standard Arabic. Ensure the translation is accurate, culturally appropriate, and suitable for an Arabic-speaking academic audience in the Middle East. Maintain the formal tone and academic style of the original text, ensuring that key technical terms, particularly those related to urban planning and green infrastructure, are translated consistently and precisely. Restructure the syntax where necessary to ensure clarity and a natural flow in Arabic, avoiding overly literal translations. Focus on delivering a coherent and well-structured abstract that adheres to the conventions of Arabic academic writing.
Text 4: Media	Translate the following text into Arabic	As a translator specializing in English-to-Arabic news translation, translate the following news article from BBC into Modern Standard Arabic. Ensure the translation is accurate, culturally appropriate, and suitable for an Arabic-speaking audience in the Middle East. Maintain the objective tone and journalistic style of the original text, ensuring consistency in key terms, particularly names and places, while restructuring the syntax where necessary for clarity and natural flow in Arabic. Avoid overly literal translations, focusing on delivering a smooth and engaging news piece that adheres to Arabic journalistic standards.

5. Results

The ratings given by the evaluators are summarized in Table 2 and shown visually in Figure 1 (Appendix 2). Overall, human translation consistently achieves the best scores across all text types. ChatGPT with detailed prompts performs very well, almost reaching the level of human translation, especially for literary and media texts. ChatGPT with simple prompts performs moderately, generally doing better than Google Translate, particularly in general and media texts. Google Translate has the weakest performance overall, struggling with fluency, cultural suitability, and style, especially in scientific and general texts.

General text

Google Translate scores the lowest with an average of 3.12, performing well in accuracy and fluency but struggling with style and cultural suitability. ChatGPT with a simple prompt scores 3.4, showing improvements in fluency and accuracy. With a detailed prompt, ChatGPT scores 3.9, excelling in accuracy and terminology. Human translation is the best, with an average of 4.46, excelling in fluency, style, and cultural suitability.

Literary text

Google Translate achieves an average of 3.66, with good results in fluency and style, but it struggles with cultural context. ChatGPT with simple prompts scores 3.7, showing good fluency and style, but it still has some difficulties with cultural elements. ChatGPT with detailed prompts scores 4.1 on average, with strong performance in accuracy and fluency. Human translation is the best, scoring 4.3, especially excelling in culture and terminology.

Table 2. Averages of the evaluators' ratings based on the text type, translation method, and evaluation dimensions

Text type	Translation methods	Evaluation dimensions					Average
		Accuracy	Fluency	Style & Tone	Cultural suitability	Terminology	
General	Google Translate	3.2	3.3	2.8	3.2	3.1	3.12
	ChatGPT -S	3.5	3.6	3.4	3.2	3.3	3.4
	ChatGPT-D	4	4	3.7	3.8	4	3.9
	Human translation	4.4	4.6	4	4.8	4.5	4.46
Literary	Google Translate	3.8	3.9	3.9	3.3	3.4	3.66
	ChatGPT -S	3.9	3.8	4	3.4	3.4	3.7
	ChatGPT-D	4.1	4.3	4.2	4	3.9	4.1
	Human translation	4.3	4.5	4.1	4.4	4.2	4.3
Scientific	Google Translate	3	2.6	3.2	2.7	2.9	2.88
	ChatGPT -S	3.4	3.2	3.4	3.1	3.1	3.24
	ChatGPT-D	4.1	4.3	3.7	4	3.6	3.94
	Human translation	4.3	4.6	3.7	4.5	4.1	4.24
Media	Google Translate	4	3.9	3.7	3.1	3.3	3.6
	ChatGPT -S	4.2	4.2	3.9	3.9	3.8	4
	ChatGPT-D	4.3	4.3	3.8	3.8	3.9	4.02
	Human translation	4.6	4.8	3.6	4.7	4.4	4.42

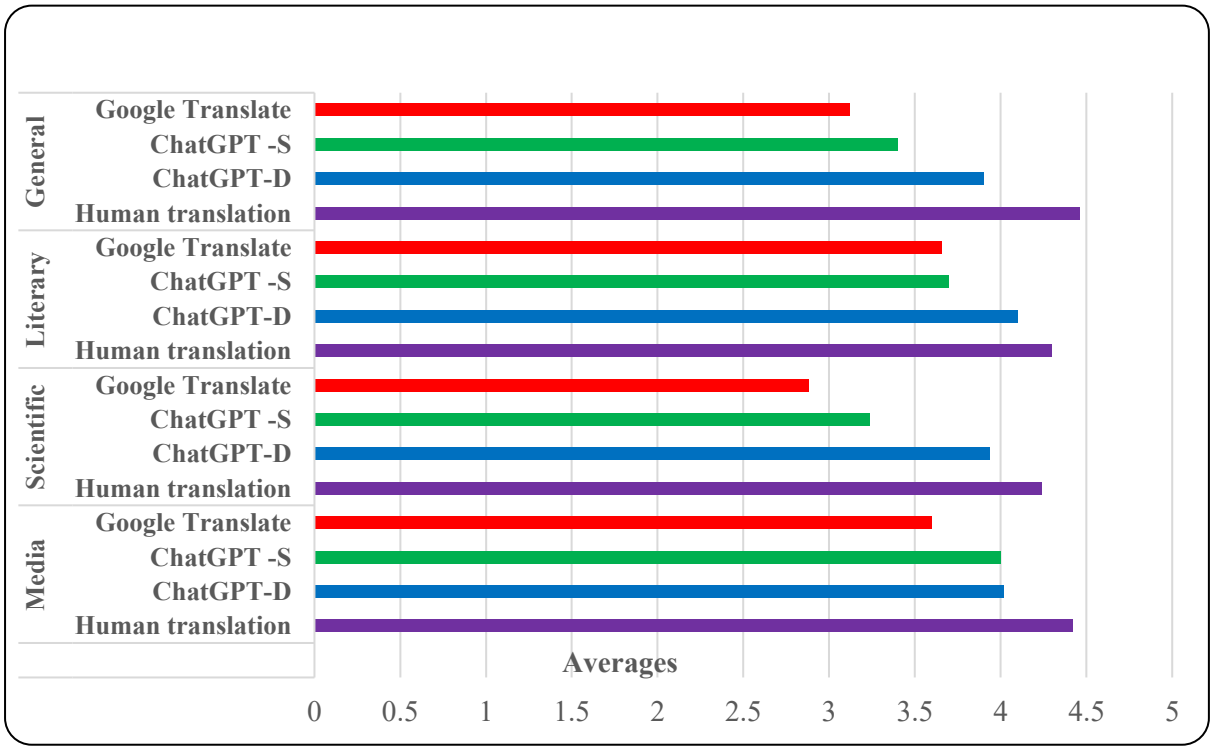

Figure 1. Averages of the evaluators' ratings based on text type and translation method

Scientific text

Google Translate scores the lowest, with an average of 2.88, and it struggles particularly with fluency and cultural suitability. ChatGPT with simple prompts performs better, with an average of 3.24, while ChatGPT with detailed prompts scores much higher, averaging 3.94, performing especially well in fluency and accuracy. Human translation has the highest score, averaging 4.24, excelling in fluency and cultural suitability.

Media text

Google Translate scores 3.6, with strong fluency and accuracy but struggling with cultural suitability. ChatGPT with simple prompts scores 4 on average, excelling in fluency and accuracy. ChatGPT with detailed prompts performs slightly better, scoring 4.02 on average, and showing consistent performance across all areas. Human translation scores the highest with an average of 4.42, especially excelling in fluency and cultural suitability.

6. Discussion

6.1. Google Translate versus ChatGPT

Google Translate relies on a sequence-to-sequence NMT model (Martin, 2024), which may be less effective at capturing subtle meanings and context. In contrast, ChatGPT, as an LLM model, was trained on a wide range of texts, enabling it to handle diverse language uses (e.g., slang, idiomatic expressions, and complex sentence structures) and giving it an advantage in understanding context (Roumeliotis and Tselikas, 2023). ChatGPT also processes texts more holistically using its generative GPT architecture, which provides greater flexibility and accuracy, particularly in maintaining context across longer or more complex texts (Chen et al., 2024). ChatGPT also demonstrates stronger cultural suitability, enabling it to better interpret idiomatic expressions, jokes, and cultural references (Abu-Rayyash, 2024). Table 3 provides an example from the general text, alongside translations by Google Translate and ChatGPT, using both simple and detailed prompts.

Table 3. An example from the translations of the general text by Google Translate and ChatGPT

Source Text	Google Translate	ChatGPT-S	ChatGPT-D
It was enough for "these men" to appear as they have in newspapers and on television for American viewers to feel a combination of anger, resentment, and fear.	وكان ظهور 'هؤلاء الرجال' كما ظهروا في الصحف وعلى شاشات التلفزيون كافياً لكي يشعر المشاهدون الأميركيون بمزيج من الغضب والاستياء والخوف.	وكان كافيًا أن يظهر 'هؤلاء الرجال' كما ظهروا في الصحف والتلفزيون ليشعر المشاهد الأمريكي بمزيج من الغضب والاستياء والخوف.	كان يكفي أن يظهر 'هؤلاء الرجال' كما شاهدهم الجمهور الأمريكي في الصحف والتلفزيون ليشعروا بمزيج من الغضب، الاستياء، والخوف.

As with the Google Translate version, the structure of the كان sentence is preserved, beginning with كان followed by its subject ظهور هؤلاء الرجال, and then its predicate كافياً. However, attaching the subordinate clause كما ظهروا في الصحف وعلى شاشات التلفزيون to the subject makes the entire structure awkward. This is especially true with the redundant word شاشات, which is unnecessary about television. Compared to this awkward structure, the two translations by ChatGPT are much smoother while effectively conveying the same message. In particular, the third translation, generated using a detailed prompt designed to produce a fluid, engaging text in a

culturally appropriate style (see Table 1), employs the modulation strategy by foregrounding "American viewers." This adjustment enhances the cohesion of the translation, making it more natural and aligned with Arabic stylistic norms.

6.2. ChatGPT: Simple prompts vs. detailed prompts

Using detailed prompts with ChatGPT has shown significant improvements compared to simpler prompts, highlighting how the quality of input affects the translation results (Gu, 2023). Providing more specific instructions helps the model meet translation goals more effectively, including considering cultural and stylistic details (Gao et al. 2023). For texts that require precision, such as scientific and literary works, this approach allows ChatGPT to manage terminology and tone better, improving the translation's accuracy and making it closer to human-level performance (Delta, 2024). This is shown by the higher ratings given to the translations created with detailed prompts (see Table 1), compared to those generated using the simple prompt "Translate the following text into Arabic." Two examples from the scientific text showing the better performance of ChatGPT using a detailed prompt versus a simple one are given in Table 4.

Table 4. Examples from ChatGPT's translations of the scientific text using a simple prompt and a detailed prompt

Source Text	ChatGPT-S	ChatGPT-D
This lack of knowledge costs us: we risk creating GSI that people find unappealing	هذا النقص في المعرفة يكلفنا: هناك خطر في إنشاء بنية تحتية قد يجدها الناس غير جذابة.	هذا النقص في المعرفة يسبب لنا بعض الخسائر: فقد نخاطر بإنشاء بنية تحتية خضراء يجدها الناس غير جذابة.
We found that landscapes with any GSI are significantly more preferred than landscapes with no GSI	وجدنا أن المناظر الطبيعية التي تحتوي على أي بنية تحتية خضراء أكثر تفضيلاً بشكل ملحوظ من تلك التي لا تحتوي على أي بنية تحتية خضراء	وجدنا أن المناظر الطبيعية التي تحتوي على أي نوع من البنية التحتية الخضراء مفضلة بشكل ملحوظ عن تلك التي لا تحتوي على أي منها

In the first example, translating "costs us" as يكلفنا through the simple prompt is vague, as the Arabic verb is left undefined, without specifying the type of cost. This could imply a financial cost, the basic meaning of the verb, but the actual meaning here refers to the risk associated with creating Green Stormwater Infrastructure (GS) that may not appeal to some people. This distinction is clarified in the translation generated with the detailed prompt as يسبب لنا بعض الخسائر. Also, the simple-prompt translation moves directly from يكلفنا to the following statement هناك خطر في إنشاء بنية تحتية without clear coherence. In contrast, the detailed-prompt translation achieves coherence by using the resultative فـ, linking the risk of "creating GSI that people find unappealing" to "this lack of knowledge."

In the second example, the phrase بنية تحتية خضراء is unnecessarily repeated twice in the simple-prompt translation, reflecting the repeated abbreviation GSI in the source text. As one human rater noted, this repetition is "awkward stylistically, particularly due to the repetition of التي لا تحتوي على بنية تحتية خضراء twice." In contrast, the detailed-prompt translation avoids redundancy by using البنية التحتية الخضراء once and addressing the repetition in the source text with the pronoun ها attached to the preposition من.

6.3. Humans outperform machines

The results above clearly show that MT has made significant progress in handling various text types across all evaluation dimensions. However, human translation still receives the highest ratings and outperforms the three MT methods (Google Translate, ChatGPT with a simple prompt, and ChatGPT with a detailed prompt) in all evaluation dimensions. Notably, the greatest differences between human translation and the MT methods are in the two dimensions that give any translation its distinct character: cultural suitability and fluency. A few clarifying examples from the literary text are given in Table 5.

Table 5. Comparison of Google Translate, ChatGPT (using simple and detailed prompts), and human translations of cultural references in the literary text

Source Text	Google Translate	ChatGPT-S	ChatGPT-D	Human
We were planning on sailing about two hundred miles a day	كنا نخطط للإبحار حوالي مائتي ميل في اليوم	كنا نخطط للإبحار حوالي مائتي ميل في اليوم	كنا نخطط للإبحار حوالي مائتي ميل في اليوم	كنا نعتزم أن نقطع مسافة 320 كيلو مترًا في اليوم
Barnacle Bill warned us about the Bay of Biscay, so we were expecting it to be bad, and it was.	حذرنا بارناكل بيل من خليج بسكاي، لذلك توقعنا أن يكون سيئًا، وكان كذلك.	حذرنا بارنكل بيل من خليج بسكاي، لذا كنا نتوقع أن يكون سيئًا، وكان كذلك.	حذرنا بارنكل بيل من خليج بسكاي، لذا كنا نتوقع أن يكون الوضع سيئًا، وقد كان كذلك.	كان بارناكل بيل قد حذرنا من خليج بسكاي، ما بين فرنسا وإسبانيا، وهكذا توقعنا سوء الأحوال الجوية فيه، وصَدَقَتْ توقعاتُنا.

Cultural factors influence the translation of measurements because they often involve adjusting to the measurement systems and conventions of the target culture. For instance, when translating imperial measurement terms (such as *inches*, *feet*, or *miles*), converting them into metric measurements (like *centimeters*, *meters*, or *kilometers*) in cultures that predominantly use the metric system may be necessary. This adaptation ensures that the translation is understandable and relevant to the target audience, aligning with their familiar cultural norms. In translating "two hundred miles" in the literary text, all MT methods used مائتي ميل, retaining the imperial measurement even though all Arab countries use the metric system (Buchholz, 2019). The only translation that converts the imperial measurement of miles to the metric measurement of kilometers is the human translation, which renders 200 miles as 320 kilometers (320 كيلو متراً).

Geographical references or names are also culturally significant in translation and require clarity, especially if the target audience is unfamiliar with them (Sadiq, 2008). Despite the large size of the Bay of Biscay (223,000 square km, according to Encyclopedia Britannica, 2024), it is not widely recognized by general Arab readers. As such, transliteration alone, such as خليج بسكاي, is insufficient, as is done in the translations provided by the three MT methods. Only the human translation includes both the transliteration and the brief explanatory phrase ما بين فرنسا وإسبانيا, offering essential context for Arab readers unfamiliar with the bay.

Fluency refers to "the extent to which the linguistic and textual expression of the target text matches the linguistic norms and conventions of the target language (grammar, spelling, cohesion, etc.) and the textual requirements of the purpose of the translation" (Lommel et al., 2015, as cited in Salmi, 2020, p. 150). The following example from the general text demonstrates how fluency is achieved by closely adhering to Arabic linguistic norms and conventions, resulting in a cohesive and natural translation.

Table 6. An example comparing fluency issues in the translations of the general text by Google Translate, ChatGPT (using simple and detailed prompts), and human translation

Source Text	Google Translate	ChatGPT-S	ChatGPT-D	Human
In order to make a point about alternative energy sources for Americans, Consolidated Edison of New York (Con Ed) ran a striking television advertisement in the summer of 1980.Film clips of various immediately recognizable OPEC personalities—Yamani, Qaddafi, and lesser-known robed Arab figures—alternated with stills and clips of other people associated with oil and Islam: Khomeini, Arafat, and Hafez al-Assad.	من أجل التأكيد على أهمية مصادر الطاقة البديلة للأميركيين، قامت شركة كونسوليديتد إيديسون في نيويورك (كون إد) ببث إعلان تلفزيوني مذهل في صيف عام 1980.وقد تناوبت مقاطع من أفلام شخصيات منظمة أوبك المعروفة على الفور. اليماني والقذافي وشخصيات عربية أقل شهرة . مع صور ومقاطع من أشخاص آخرين مرتبطين بالنفط والإسلام: الخميني وعرفات وحافظ الأسد.	من أجل توضيح نقطة حول مصادر الطاقة البديلة للأمريكيين، نشرت شركة كون إيديسون في نيويورك (كون إد) إعلانًا تلفزيونيًا ملفتًا في صيف عام 1980. ظهرت في الإعلان مقاطع لأشخاص من منظمة أوبك معروفين فورًا مثل يماني، القذافي، وشخصيات عربية أخرى أقل شهرة ترتدي العباءات، بالتناوب مع لقطات ثابتة ومتحركة لأشخاص آخرين مرتبطين بالنفط والإسلام: الخميني، عرفات، وحافظ الأسد.	في صيف عام 1980، عرضت شركة كونسوليديتد إديسون في نيويورك (كون إد) إعلانًا تلفزيونيًا لافتًا للنظر بهدف تسليط الضوء على مصادر الطاقة البديلة للأمريكيين. تضمنت اللقطات شخصيات بارزة من منظمة أوبك، مثل يماني، القذافي، وشخصيات عربية أخرى مرتدية الجلباب أقل شهرة، بالتناوب مع صور ثابتة ومقاطع لأشخاص آخرين مرتبطين بالنفط والإسلام: الخميني، عرفات، وحافظ الأسد.	عندما أرادت شركة إديسون المتحدة بنيويورك (شركة كون إيد) أن تُقْنع الأمريكيين بضرورة توفير مصادرَ بديلةٍ للطاقة، أذاعت إعلانًا تلفزيونيًّا مثيرًا في صيف عام 1980م، يتضمن لقطاتٍ متحرِّكةً قديمة لبعض الشخصيات المعروفة في منظمة البلدان المُصدِّرة للنفط (أوبك) — مثل الدكتور أحمد زكي يماني، والعقيد معمَّر القذافي، وبعض الشخصيات العربية التي تلبس الزي العربي وإن تكن أقلَّ شهرة — ويمزج بينها، بالتناوب، وبين بعض اللقطات الثابتة الأخرى، إلى جانب لقطاتٍ لشخصيات أخرى ارتبطتْ أسماؤها بالنفط والإسلام مثل الخوميني، وعرفات، وحافظ الأسد.

The human translation is the only one to merge two source text sentences into a single cohesive sentence in Arabic, ensuring a smooth flow between ideas. This approach maintains fluency by effectively using proper punctuation, making the long sentence clear and easy to follow. Transitional words, such as مثل, have been added more frequently in the human translation (e.g., twice before the personalities' names) compared to only once in ChatGPT's translations, enhancing the natural progression of the text. Also, the human translation consistently uses the Arabic addition letter و along with commas to separate successive nouns, aligning with Arabic linguistic norms. In contrast, the other MT methods mostly apply the English style, using و only once before the last noun.

To enhance clarity, the human translation includes additional names and titles absent in the source text. For example, الدكتور أحمد زكي[5] is added before يماني to avoid any confusion with *Yemen* or *Yemeni*. Also, while the three MT methods simply transliterate "Consolidated," the human translation renders it as المتحدة, making it more accessible to Arab readers who may not be familiar with the English term. Since "Consolidated" is not commonly used in transliteration, it would be difficult for Arab readers to recognize or understand. Although OPEC is widely known among Arab readers, the human translation follows the prevalent Arabic style by paraphrasing the acronym as منظمة البلدان المُصدِّرة للنفط (أوبك) and providing its

[5] Ahmed Zaki Yamani (1930–2021) was the Minister of Petroleum and Mineral Resources in Saudi Arabia (1962–1986) and represented the country in OPEC for 25 years.

transliteration. This dual approach enhances clarity and aligns with Arabic conventions.

6.4. Challenges and opportunities in using MT

MT is expected to play an increasingly significant role in translation, particularly for languages like Arabic, which remain underrepresented in global MT research. These tools can enhance the speed and cost-effectiveness of translation, especially for informal tasks or projects with tight deadlines. Integrating various MT tools—such as NMT and LLMs—can help professionals manage large volumes of work more efficiently. However, human expertise remains essential for translating complex or culturally sensitive content. The use of MT tools in critical domains such as law and medicine also raises ethical concerns, as they may struggle with nuanced terminology, potentially leading to misunderstandings or harmful consequences. As AI continues to advance, the role of human translators will be crucial in preserving linguistic diversity and providing the contextual and cultural insights that machines cannot replicate. While MT tools can support the translation process, human intervention remains indispensable for tasks requiring deep cultural understanding and ethical judgment.

7. Conclusion

This study compared NMT, represented by Google Translate, and LLMs, represented by ChatGPT, with human translation, focusing on their strengths, weaknesses, and practical applications. The results show that while MT has advanced, human translation remains superior in cultural suitability, fluency, and accuracy across various text types. ChatGPT, with detailed prompts, performed nearly as well as human translation, especially in literary and media texts. At the same time, Google Translate struggled with fluency, cultural suitability, and terminology, particularly in scientific and general texts. The study highlights the importance of prompt quality in improving MT output, but human translation remains unmatched in overall quality, underscoring its crucial role in producing precise, culturally adjusted translations.

This study offers valuable insights for translation trainers, professionals, and students. For trainers, it provides a basis for developing curricula that address both the strengths and limitations of MT and human translation, equipping students to use these tools effectively while emphasizing the importance of human expertise. For translators, the study offers guidance on selecting tools based on text type and task requirements, highlighting when human translation is essential and when MT can serve as a time-saving aid requiring post-editing. Students benefit from understanding the interplay between MT and human translation, developing critical thinking skills, and recognizing the unique value of human judgment in complex or specialized translation tasks.

Despite its contributions, this study has some limitations. It examined a limited range of text types—general, literary, scientific, and media—which may not cover all translation challenges. Also, it employed a narrow set of evaluation dimensions, potentially missing other important aspects of translation quality, and relied on a small number of human evaluators, which could have introduced bias or inconsistency. Also, the focus on translations from English into Arabic makes it harder to apply the results to Arabic-to-English translations, which might have different challenges.

Future research could address these gaps by examining more specialized or creative text types and using a broader range of evaluation dimensions to evaluate translation quality. Increasing the number of human evaluators would enhance the reliability of evaluations. Comparative studies of English-to-Arabic and Arabic-to-English translations could reveal unique challenges and insights related to directionality. Moreover, exploring the cognitive processes involved in post-editing MT outputs and how human intervention improves translation quality could provide deeper insights into the interaction between MT systems and human translators.

Acknowledgment:
I sincerely thank all the evaluators who generously dedicated their time and expertise to evaluating the translations investigated. Your valuable feedback has profoundly influenced the outcomes of this study. I also extend my heartfelt gratitude to Prof. Christine Royce of Shippensburg University, USA, for her assistance in providing materials and support.

References

Abu-Rayyash, H. (2024). AI meets comedy: Viewers' reactions to GPT-4 generated humor translation. *Ampersand, 12*, 1-8. https://doi.org/10.1016/j.amper.2023.100162

Aiken, M. (2019). An updated evaluation of Google Translate accuracy. *Studies in Linguistics and Literature, 3*(3), 253-260. https://doi.org/10.22158/sll.v3n3p253

Aldawsari, H. A. (2024). Evaluating translation tools: Google Translate, Bing Translator, and Bing AI on Arabic colloquialisms. *Arab World English Journal*, 237-251. https://dx.doi.org/10.24093/awej/ChatGPT.16

Algobaei, F., Alzain, E., Naji, E., & Nagi, K. A. (2024). Gender issues between Gemini and ChatGPT: The case of English-Arabic translation. *World Journal of English Language, 15*(1), 9-16. https://doi.org/10.5430/wjel.v15n1p9

Al-Kaabi, M. H., AlQbailat, N. M., Badah, A., Ismail, I. A., & Hicham, K. B. (2024). Examining the cultural connotations in human and machine translations: A corpus study of Naguib Mahfouz's *Zuqāq al-Midaqq*. *Journal of Language Teaching and Research, 15*(3), 707-718. https://doi.org/10.17507/jltr.1503.03

Alkhawaja, L. (2024). Unveiling the new frontier: ChatGPT-3 powered translation for Arabic-English language pairs. *Theory and Practice in Language Studies, 14*(2), 347-357. https://doi.org/10.17507/tpls.1402.05

Al-khresheh, M. H., & Almaaytah, S. A. (2018). English proverbs into Arabic through Machine Translation. *International Journal of English Language & Translation Studies, 7*(5), 15-19.

Alowedi, N. A., & Al-Ahdal, A. A. M. H. (2023). Artificial Intelligence based Arabic-to-English machine versus human translation of poetry: An analytical study of outcomes. *Journal of Namibian Studies: History Politics Culture, 33*, 1523-1538. https://doi.org/10.59670/jns.v33i.800

Al-Qinai, J. (2000). Translation quality assessment: Strategies, parameters and procedures. *Meta, 45*(3), 497–519. https://doi.org/10.7202/001878ar

Al-Sabbagh, R. (2024). The negative transfer effect on the Neural Machine Translation of Egyptian Arabic adjuncts into English: The case of Google Translate. *International Journal of Arabic-English Studies (IJAES), 24*(1), 95-118. https://doi.org/10.33806/ijaes.v24i1.560

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural Machine Translation by jointly learning to align and translate. In Y. Bengio & Y. LeCun (Eds.), *Proceedings of the 3rd international conference on learning representations (ICLR'15)* (pp. 1-15). https://arxiv.org/abs/1409.0473

Banimelhem, O., & Amayreh, W. (2023). Is ChatGPT a good English to Arabic Machine Translation tool? In *2023 14th International Conference on Information and Communication Systems (ICICS)* (pp. 1-6). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICICS60529.2023.10330525

Hatim, B., & Mason, I. (1990). *Discourse and the translator*. Longman.

Buchholz, K. (2019, June 6). Metric or imperial? *Statista Daily Data*. https://www.statista.com/chart/18300/countries-using-the-metric-or-the-imperial-system/

Caswell, I. (2024, June 27). 110 new languages are coming to Google Translate. *Google Translate Blog*. https://blog.google/products/translate/google-translate-new-languages-2024/

Chen, J., Wang, S., Qi, Z., Zhang, Z., Wang, C., & Zheng, H. (2024). A combined encoder and transformer approach for coherent and high-quality text generation. *arXiv.org*. https://arxiv.org/abs/2411.12157

Daniele, F. (2019). Performance of an automatic translator in translating medical abstracts. *Heliyon, 5*(10), 1-6. https://doi.org/10.1016/j.heliyon.2019.e02687

Delta, R. (2024, January 6). Use this guide to master ChatGPT in 6 simple steps. *Medium*. https://medium.com/@msgforrobin/use-this-guide-to-master-chatgpt-in-6-simple-steps-f7e3869a8bf

Ducar, C., & Schocket, D. H. (2018). Machine Translation and the L2 classroom: Pedagogical solutions for making peace with Google Translate. *Foreign Language Annals, 51*(4), 779-795. https://doi.org/10.1111/flan.12366

El-Saadany, M. (2024). A comparative study between ChatGPT and human translation in translating English proverbs into Arabic. *Journal of Scientific Research in Arts, 25*(5), 24-54. https://doi.org/10.21608/jssa.2024.257874.1592

Encyclopedia Britannica. (2024, March 13). Bay of Biscay. https://www.britannica.com/place/Bay-of-Biscay

Gánem, A. (2024, February 20). How to use ChatGPT for translations: Updated guide 2024. https://www.websiteplanet.com/blog/how-to-use-chatgpt-for-translation/

Gao, Y., Wang, R., & Hou, F. (2023). How to design translation prompts for ChatGPT: An empirical study. *arXiv.org*. https://doi.org/10.48550/arXiv.2304.02182

Giray, L. (2023). Prompt engineering with ChatGPT: A guide for academic writers. *Annals of Biomedical Engineering, 51*, 2629–2633. https://doi.org/10.1007/s10439-023-03272-4

Gu, W. (2023). Linguistically informed ChatGPT prompts to enhance Japanese-Chinese machine translation: A case study on attributive clauses. *arXiv.org*. https://doi.org/10.48550/arXiv.2303.15587

Hearne, M., & Way, A. (2011). Statistical Machine Translation: A guide for linguists and translators. *Language and Linguistics Compass, 5*(5), 205-226.

House, J. (1981). *A model for translation quality assessment*. TBL-Verlag Narr.

Hutchins, W. J. (2005). Example-Based Machine Translation: A review and commentary. *Machine Translation, 19*(3-4), 197-211. https://doi.org/10.1007/s10590-006-9003-9

Hutchins, W. J., & Somers, H. L. (1992). *An introduction to Machine Translation*. Academic Press.

Jiang, Z., & Zhang, Z. (2024). Can ChatGPT rival Neural Machine Translation? A comparative study. *arXiv.org*. https://arxiv.org/html/2401.05176v1

Jiao, W., Wang, W., Huang, J.-T., Wang, X., Shi, S., & Tu, Z. (2023). Is ChatGPT a good translator? Yes with GPT-4 as the engine. *arXiv.org*. https://doi.org/10.48550/arXiv.2301.08745

Khoong, E. C., Steinbrook, E., Brown, C., & Fernandez, A. (2019). Assessing the use of Google Translate for Spanish and Chinese translations of emergency department discharge instructions. *JAMA Internal Medicine, 179*(4), 580-582. https://doi.org/10.1001/jamainternmed.2018.7653

Martin, S. (2024). Advancements in Neural Machine Translation: Techniques and applications. *Journal of Innovative Technologies, 7*(1), 1-9.

Mohamed, A. H. (2023). Google Translate: Lexical and syntactic problems of FIDIC translation into Arabic. *CDELT Occasional Papers in the Development of English Education, 83*(1), 33-53. https://dx.doi.org/10.21608/opde.2023.325327

Moneus, A. M., & Sahari, Y. (2024). Artificial Intelligence and human translation: A contrastive study based on legal texts. *Heliyon, 10*(6), 1-12. https://doi.org/10.1016/j.heliyon.2024.e28106

Morpurgo, M. (1999). *Kensuke's Kingdom*. Farshore.

Morpurgo, M. (2023). *Mamlakit Kinsūki* (M. Enani, Trans.). Hindawi. https://www.hindawi.org/books/46315048/ (Original work published 1999)

Moslem, Y., Haque, R., Kelleher, J. D., & Way, A. (2023). Adaptive machine translation with large language models. *arXiv.org*. https://doi.org/10.48550/arXiv.2301.13294

Nagi, K. A. (2023). Arabic and English relative clauses and Machine Translation challenges. *Journal of Social Studies, 29*(3), 145-165. https://doi.org/10.20428/jss.v29i3.2180

Nagi, K. A., Alzain, E., & Naji, E. (2024). Informed prompts and improving ChatGPT English to Arabic translation. *Al-Andalus Journal for Humanities & Social Sciences, 11*(98), 211-228. https://doi.org/10.35781/1637-000-098-007

Newmark, P. (1988). *A textbook of translation*. Prentice-Hall International.

Okpor, M. D. (2014). Machine Translation approaches: Issues and challenges. *International Journal of Computer Science Issues (IJCSI), 11*(5), 159-165.

Papineni, K., Roukos, S., Todd, W., Henderson, J., & Reeder, F. (2022). Corpus-based comprehensive and diagnostic MT evaluation: Initial Arabic, Chinese, French, and Spanish results. In M. Marcus (Ed.), *HLT '02: Proceedings of the Second International Conference on Human Language Technology Research* (pp. 132-137). Morgan Kaufmann Publishers.

Popović, M. (2015). chrF: Character N-Gram F-Score for automatic MT evaluation. *Proceedings of the Tenth Workshop on Statistical Machine Translation*, 392-395.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are unsupervised multitask learners. *OpenAI Blog, 1*(8). https://insightcivic.s3.us-east-1.amazonaws.com/language-models.pdf

Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and Open-AI models: A preliminary review. *Future Internet, 15*(6), 1-24. https://doi.org/10.3390/fi15060192

Sadiq, S. (2008). Translation: Some lexical and syntactic problems & suggested solutions. *Wata Translation Journal, 5*(1), 1-22.

Said, E. W. (1981). *Covering Islam: How the media and the experts determine how we see the rest of the world*. Vintage.

Said, E. W. (2022). *Taghṭiyat 'al-'Islām: Kayfa tataḥakkam 'ajhizat 'al-'i'lām wa yatahakkam al-khubarā' fi ru'yatina li-sā'ir buldān 'al-'ālam* (M. Enani, Trans.). Hindawi. https://www.hindawi.org/books/18028482/ (Original work published 1981)

Salmi, L. (2020). Fluency in evaluating and assessing translations. In P. Lintunen, M. Mutta, & P. Peltonen (Eds.), *Fluency in L2 learning and use* (pp. 146-165). Language Matters.

Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makho, J. (2006). A study of translation error rate with targeted human annotation. *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas (AMTA 06)* (pp. 223–231). Association for Machine Translation in the Americas.

Son, J., & Kim, B. (2023). Translation performance from the user's perspective of Large Language Models and Neural Machine Translation systems. *Information, 14*(10), 1-18. https://doi.org/10.3390/info14100574

Suppakittpaisarn, P., Larsen, L., & Sullivan, W. C. (2019). Preferences for green infrastructure and green stormwater infrastructure in urban landscapes: Differences between designers and laypeople. *Urban Forestry & Urban Greening, 43*(1), 1-7. https://doi.org/10.1016/j.ufug.2019.126378

Taira, B. R., Kreger, V., Orue, A., & Diamond, L. C. (2021). A pragmatic assessment of Google Translate for emergency department instructions. *Journal of General Internal Medicine, 36*(11), 3361-3365. https://doi.org/10.1007/s11606-021-06666-z

Taleghani, M., & Pazouki, E. (2018). Free online translators: A comparative assessment in terms of idioms and phrasal verbs. *International Journal of English Language and Translation Studies, 6*(1), 15-19.

Yousif, N. (2024a, August 9). ABC News: Trump-Harris agree to presidential debate. *BBC*. https://www.bbc.com/news/articles/cy54l4x7d71o

Yousif, N. (2024b, August 9). Munāẓara bayna Tramb wa Haris fī Sibtambir 'al-muqbil. *bī bī sī*. https://www.bbc.com/arabic/articles/cdxlv2n8ekro

Zong, Z. (2018). Research on the relations between Machine Translation and human translation. *Journal of Physics: Conference Series, 1087*(6), 1-7. https://doi.org/10.1088/1742-6596/1087/6/062046

Appendix 1: Translation Evaluation Questionnaire

Welcome to My Survey!

Thank you for taking the time to participate! We are currently evaluating four different translations of various texts, and your feedback is essential for this project.

In this survey, you will find a selection of original texts paired with their translations. We encourage you to carefully review each translation, considering its accuracy, style, and overall readability. Your thoughtful evaluation will provide valuable insights into how effectively these translations convey the original meaning and tone.

Once you have reviewed the texts, please complete the questionnaire based on the questions provided. Your opinions and observations will greatly assist us in understanding the strengths and weaknesses of each translation.

Thank you again for your participation!

For any inquiries, please contact me at: saudi.sadiq@mu.edu.eg

==========

Section 1: Background Information

Name (Optional):

Institution/University (Optional):

Current Position:

Years of Experience in Translation (Give the number):

Section 2: Evaluation of Text 1 Translations: Please review the following text and its four Arabic translations. Then, evaluate each based on the questions provided below.

Source Text:

Islam and the West

In order to make a point about alternative energy sources for Americans, Consolidated Edison of New York (Con Ed) ran a striking television advertisement in the summer of 1980. Film clips of various immediately recognizable OPEC personalities—Yamani, Qaddafi, and lesser-known robed Arab figures—alternated with stills and clips of other people associated with oil and Islam: Khomeini, Arafat, and Hafez al-Assad. None of these figures was mentioned by name, but we were told ominously that "these men" control America's sources of oil. The solemn voice-over in the background made no reference to who "these men" actually are or where they come from, leaving it to be felt that this all-male cast of villains has placed Americans in the grip of an unrestrained threat. It was enough for "these men" to appear as they have in newspapers and on television for American viewers to feel a combination of anger, resentment, and fear. This combination of feelings was what Con Ed sought to arouse and exploit for domestic commercial reasons, just as a year earlier, Stuart Eizenstat, President Carter's domestic policy adviser and now a senior official in the Clinton administration, had urged the president to "with strong steps, mobilize the nation around a real crisis and with a clear enemy—OPEC."

Translation 1:

الإسلام والغرب

من أجل التأكيد على أهمية مصادر الطاقة البديلة للأميركيين، قامت شركة كونسوليديتد إيديسون في نيويورك (كون إد) ببث إعلان تلفزيوني مذهل في صيف عام 1980. وقد تناوبت مقاطع من أفلام شخصيات منظمة أوبك المعروفة على الفور. اليماني والقذافي وشخصيات عربية أقل شهرة. مع صور ومقاطع من أشخاص آخرين مرتبطين بالنفط والإسلام: الخميني وعرفات وحافظ الأسد.

ولم يذكر الإعلان أياً من هذه الشخصيات بالاسم، ولكننا سمعنا على نحو ينذر بالسوء أن "هؤلاء الرجال" يسيطرون على مصادر النفط في أميركا. ولم يشر الصوت المهيب في الخلفية إلى هوية "هؤلاء الرجال" أو من أين أتوا، الأمر الذي جعلنا نشعر بأن هذه المجموعة من الأشرار الذكور قد وضعت الأميركيين في قبضة تهديد غير مقيد. وكان ظهور "هؤلاء الرجال" كما ظهروا في الصحف وعلى شاشات التلفزيون كافياً لكي يشعر المشاهدون الأميركيون بمزيج من الغضب والاستياء والخوف. وكان هذا المزيج من المشاعر هو ما سعت شركة كون إيد إلى إثارته واستغلاله لأسباب تجارية محلية، تماماً كما حث ستيوارت إيزنستات، مستشار الرئيس كارتر للسياسة الداخلية والذي أصبح الآن مسؤولاً كبيراً في إدارة كلينتون، الرئيس قبل عام واحد على "حشد الأمة حول أزمة حقيقية ومع عدو واضح . أوبك"، من خلال "خطوات قوية".

Translation 2:

الإسلام والغرب

من أجل توضيح نقطة حول مصادر الطاقة البديلة للأمريكيين، نشرت شركة كون إيديسون في نيويورك (كون إد) إعلانًا تلفزيونيًا ملفتًا في صيف عام 1980. ظهرت في الإعلان مقاطع لأشخاص من منظمة أوبك معروفين فورًا مثل يماني، القذافي، وشخصيات عربية أخرى أقل شهرة ترتدي العباءات، بالتناوب مع لقطات ثابتة ومتحركة لأشخاص آخرين مرتبطين بالنفط والإسلام: الخميني، عرفات، وحافظ الأسد. لم يُذكر أي من هؤلاء الأشخاص بأسمائهم، لكن قيل لنا بتحذير أن "هؤلاء الرجال" يتحكمون في مصادر النفط الأمريكية. الصوت الخلفي الجاد لم يذكر من هم "هؤلاء الرجال" أو من أين جاءوا، مما جعل الأمر يبدو أن هذه المجموعة من الأشرار الذكور قد وضعت الأمريكيين في قبضة تهديد غير محدود. وكان كافيًا أن يظهر "هؤلاء الرجال" كما ظهروا في الصحف والتلفزيون ليشعر المشاهد الأمريكي بمزيج من الغضب والاستياء والخوف. هذا المزيج من المشاعر كان ما سعت شركة كون إيد إلى إثارته واستغلاله لأسباب تجارية محلية، تمامًا كما في العام السابق، حيث كان ستيوارت إيزنستات، مستشار السياسة الداخلية للرئيس كارتر والآن مسؤول رفيع في إدارة كلينتون، قد حث الرئيس على "اتخاذ خطوات قوية، وتعبئة الأمة حول أزمة حقيقية وعدو واضح—أوبك."

Translation 3:

الإسلام والغرب

في صيف عام 1980، عرضت شركة كونسوليديتد إديسون في نيويورك (كون إد) إعلانًا تلفزيونيًا لافتًا للنظر بهدف تسليط الضوء على مصادر الطاقة البديلة للأمريكيين. تضمنت اللقطات شخصيات بارزة من منظمة أوبك، مثل يماني، القذافي، وشخصيات عربية أخرى مرتدية الجلباب أقل شهرة، بالتناوب مع صور ثابتة ومقاطع لأشخاص آخرين مرتبطين بالنفط والإسلام: الخميني، عرفات، وحافظ الأسد. لم يتم ذكر أي من هذه الشخصيات بالاسم، لكن تم إبلاغنا بشكل مقلق أن "هؤلاء الرجال" يسيطرون على مصادر النفط الأمريكية. لم يذكر الصوت الجاد في الخلفية أي تفاصيل حول من هم "هؤلاء الرجال" أو من أين أتوا، تاركًا المشاهد يشعر بأن هذه المجموعة من الأشرار الذكور قد وضعت الأمريكيين في قبضة تهديد غير مقيد. كان يكفي أن يظهر "هؤلاء الرجال" كما شاهدهم الجمهور الأمريكي في الصحف والتلفزيون ليشعروا بمزيج من الغضب، الاستياء، والخوف. وكان هذا المزيج من المشاعر هو ما سعت شركة كونسوليديتد إديسون إلى إثارته واستغلاله لأغراض تجارية داخلية، تمامًا كما نصح ستيوارت إيزنستات، مستشار الرئيس كارتر للسياسات الداخلية والذي أصبح الآن مسؤولًا رفيعًا في إدارة كلينتون، الرئيس قبل عام بـ "اتخاذ خطوات قوية، لحشد الأمة حول أزمة حقيقية وعدو واضح—أوبك."

Translation 4:

الإسلام والغرب

عندما أرادت شركة إديسون المتحدة بنيويورك (شركة كون إيد) أن تُقْنع الأمريكيين بضرورة توفير مصادرَ بديلةٍ للطاقة، أذاعت إعلانًا تلفزيونيًّا مثيرًا في صيف عام 1980م، يتضمن لقطاتٍ متحرِّكةً قديمة لبعض الشخصيات المعروفة في منظمة البلدان المُصدِّرة للنفط (أوبك) — مثل الدكتور أحمد زكي يماني، والعقيد معمَّر القذافي، وبعض الشخصيات العربية التي تلبس الزي العربي وإن تكن أقلَّ شهرة — ويمزج بينها، بالتناوب، وبين بعض اللقطات الثابتة الأخرى، إلى جانب لقطاتٍ لشخصيات أخرى ارتبطتْ أسماؤها بالنفط والإسلام مثل الخوميني، وعرفات، وحافظ الأسد. ولم يُشِر الإعلان إلى أيٍّ من هذه الشخصيات بأسمائها، ولكن المذيع قال بصوت المنذر المحذِّر إن «هؤلاء الرجال» يتحكمون في مصادر النفط الأمريكية. وكان صوت المذيع القادم من الخلفية ذا نبراتٍ وَقورة، ولم يُفصح عن أسماء «هؤلاء الرجال» ولا عن البلدان التي ينتمون إليها، بل ترك المشاهدين يشعرون بأن هذه الكوكبة من الأشرار الذكور قد أوقَعوا الأمريكيِّين في قبضةِ من يتلذَّذ بتعذيبهم دونما ضابطٍ أو رابط. وكان يكفي أن يظهر «هؤلاء الرجال» على النحو الذي ظهروا به في الصحف والتلفزيون حتى يعتريَ الأمريكيِّين مزيجٌ من مشاعر الغضب والاستياء والخوف. وكانت هذه المشاعر هي التي عمدت شركة «كون إيد» إلى إثارتها واستغلالها فورًا لأسباب تجارية محلية، تمامًا كما حدث قبل عام واحد، عندما ألحَّ ستيوارت أيزنستات، مستشار الرئيس كارتر للسياسات المحلية، على الرئيس أن "يتخذ خطوات قوية لحشد الأمة والالتفاف حول أزمةٍ حقيقية وتحديد عدوٍّ واضح لنا؛ منظمة أوبك."

==========

Section 3: Evaluation of Text 2 Translations: Please review the following text and its four

Arabic translations. Then, evaluate each based on the questions provided below.

Source Text:

Ship's Log

September 20

It's five in the morning. I'm on watch in the cockpit and no one else is awake. We left Southampton ten days ago now. The Channel was full of tankers. There were dozens of them. So, either Mom or Dad took turns on watch the first two nights. They wouldn't let me. I don't know why not. There wasn't any fog, and I can see as well as they can.

We were planning on sailing about two hundred miles a day—that's about eight knots. But in the first week, we were lucky if we made fifty miles a day.

Barnacle Bill warned us about the Bay of Biscay, so we were expecting it to be bad, and it was. Force 9 gale. Force 10 sometimes. We were slammed all over the place. I thought we'd sink. I really did. Once, when we came to the top of a wave, I saw the bow of the Peggy Sue pointing straight up at the moon. It was like she was going to take off. Then we were hurled down the other side so fast, I was sure we were going to the bottom. It was bad. I mean, it was horrible, really horrible. But the Peggy Sue didn't fall apart, and we made it to Spain.

Translation 1:

سجل السفينة

20سبتمبر

الساعة الخامسة صباحًا. أنا أراقب في قمرة القيادة ولا أحد مستيقظ. غادرنا ساوثهامبتون منذ عشرة أيام الآن. كانت القناة مليئة بناقلات النفط. كان هناك العشرات منها. لذلك، تناوب أمي أو أبي على المراقبة في أول ليلتين. لم يسمحا لي بذلك. لا أعرف لماذا. لم يكن هناك أي ضباب، ويمكنني الرؤية جيدًا مثلهما.

كنا نخطط للإبحار حوالي مائتي ميل في اليوم - أي حوالي ثماني عقد. ولكن في الأسبوع الأول، كنا محظوظين إذا قطعنا خمسين ميلاً في اليوم.

حذرنا بارناكل بيل من خليج بسكاي، لذلك توقعنا أن يكون سيئًا، وكان كذلك. عاصفة بقوة 9. وأحيانًا بقوة 10. لقد ضربتنا الأمواج في كل مكان. اعتقدت أننا سنغرق. لقد غرقت بالفعل. ذات مرة، عندما وصلنا إلى قمة موجة، رأيت قوس سفينة بيجي سو يشير مباشرة إلى القمر. كان الأمر كما لو أنها كانت على وشك الإقلاع. ثم ألقينا إلى الجانب الآخر بسرعة كبيرة، لدرجة أنني كنت متأكدًا من أننا سنصل إلى القاع. كان الأمر سيئًا. أعني، كان الأمر مروعًا، مروعًا حقًا. لكن سفينة بيجي سو لم تتفكك، ووصلنا إلى إسبانيا.

Translation 2:

سجل السفينة

20 سبتمبر

الساعة الخامسة صباحًا. أنا في قمرة القيادة للحراسة، ولا أحد آخر مستيقظ. غادرنا ساوثهامبتون منذ عشرة أيام. كانت القناة مليئة بناقلات النفط. كانت هناك العشرات منها. لذا، تولى أبي أو أمي الحراسة في الليلتين الأوليين. لم يسمحوا لي بالقيام بذلك. لا أعرف لماذا. لم يكن هناك ضباب، وأنا أستطيع الرؤية مثلهم تمامًا.

كنا نخطط للإبحار حوالي مائتي ميل في اليوم—أي حوالي ثمانية عقد. ولكن في الأسبوع الأول، كنا محظوظين إذا قطعنا خمسين ميلًا في اليوم.

حذرنا بارنكل بيل من خليج بسكاي، لذا كنا نتوقع أن يكون سيئًا، وكان كذلك. عاصفة من الدرجة التاسعة. وأحيانًا من الدرجة العاشرة. كنا نتقلب في كل اتجاه. اعتقدت أننا سنغرق. حقًا كنت أظن ذلك. مرة، عندما وصلنا إلى قمة موجة، رأيت مقدمة بيجي سو تشير مباشرة نحو القمر. كان الأمر وكأنها ستقلع. ثم اندفعنا نحو الجانب الآخر بسرعة كبيرة، وكنت متأكدًا أننا في طريقنا إلى القاع. كان الأمر سيئًا. أعني، كان فظيعًا، حقًا فظيعًا. لكن بيجي سو لم تتحطم، ووصلنا إلى إسبانيا.

Translation 3:

سجل السفينة

20 سبتمبر

الساعة الخامسة صباحًا. أنا أقوم بالمراقبة في قمرة القيادة والجميع لا يزالون نائمين. لقد غادرنا ساوثهامبتون قبل عشرة أيام. كان القنال مليئًا بناقلات النفط. كانت هناك عشرات منها. لذا، تولى والداي المراقبة في الليلتين الأوليين. لم يسمحا لي بذلك. لا أعرف لماذا. لم يكن هناك ضباب، وأنا أستطيع الرؤية مثلهم تمامًا.

كنا نخطط للإبحار حوالي مائتي ميل في اليوم - أي حوالي ثماني عقد بحرية. لكن في الأسبوع الأول، كنا محظوظين إذا قطعنا خمسين ميلاً في اليوم.

حذرنا بارنكل بيل من خليج بسكاي، لذا كنا نتوقع أن يكون الوضع سيئًا، وقد كان كذلك. عاصفة بحرية بقوة 9، وأحيانًا تصل إلى قوة 10. كنا نتعرض للضرب من كل جانب. اعتقدت أننا سنغرق، حقًا. في مرة، عندما وصلنا إلى قمة موجة، رأيت مقدمة "بيجي سو" تشير مباشرة إلى القمر. كان الأمر وكأنها ستنطلق في السماء. ثم قذفنا بسرعة على الجانب الآخر، وكنت متأكدًا أننا سنتجه نحو القاع. كان الوضع سيئًا. أعني، كان فظيعًا، حقًا فظيعًا. لكن "بيجي سو" لم تتحطم، وتمكنا من الوصول إلى إسبانيا .

Translation 4:

سجل السفينة

20 سبتمبر

الساعة الآن الخامسة صباحًا. وأنا أقوم بنوبة المراقبة في غرفة القيادة، والجميع نائمون. تركنا ساوثامتون منذ عشرة أيام، وكان القنال الإنجليزي مليئًا بناقلات النِّفْطِ. كانت عشرات الناقلات تغدو وتروح. وهكذا كان أبي وأمي يتبادلان المراقبة في الليلتين الأوليين، ولم يسمحَا لي بذلك. لا أدري لِمَ لا، لم يكن في الجو أيُّ ضباب، وقدرتي على الرؤية لا تقل عن قدرتهما.

كنا نعتزم أن نقطع مسافة 320 كيلو مترًا في اليوم، أي أن نسير بسرعة ثماني عُقَد، ولكننا لم نستطع تجاوز 80 كيلو مترًا يوميًّا في الأسبوع الأول.

كان بارناكل بيل قد حذرنا من خليج بسكاي، ما بين فرنسا وإسبانيا، وهكذا توقعنا سوء الأحوال الجوية فيه، وصَدَقَتْ توقعاتُنا. كانت قوة الريح فيه تصل إلى 9 وأحيانًا إلى 10 عُقَد، وكانت الريح تتقاذفنا هنا وهناك. وظننت أننا سوف نغرق. بل كنت أعتقد ذلك حقًّا، وذات يوم عندما حملتنا موجة عالية رأيت مقدم السفينة بيجي سو يشير إلى أعلى، نحو القمر، فكأنما كانت سوف تنطلق إليه، وإذا بنا ننحدر إلى الجانب الآخر بسرعة خارقة حتى تصورتُ أننا سنغوص إلى القاع. كان الموقف سيئًا. أقصد أنه كان رهيبًا، رهيبًا حقًّا. ولكن بيجي سو لم تتفتت، ونجحنا في الوصول إلى إسبانيا.

==========

Section 4: Evaluation of Text 3 Translations: Please review the following text and its four Arabic translations. Then, evaluate each based on the questions provided below.

Source Text:

Many cities in the US have implemented Green Stormwater Infrastructure (GSI) to manage stormwater. Because several types of GSI are relatively new, we do not know how people perceive GSI or the extent to which they prefer various configurations of GSI. We also do not know the extent to which laypeople perceive GSI differently from designers. This lack of knowledge costs us: we risk creating GSI that people find unappealing, and if they do not like it, they may not want to spend time in or near it, they may not maintain it, and they may not reap health benefits from it. In this study, 497 laypeople and 117 designers provided preference ratings of 55 photographs of GSI. We used factor analysis to identify the categories of GSI that people perceive. We found that landscapes with any GSI are significantly more preferred than landscapes with no GSI, and that the messiness of GSI impacted preference. Designers' preferences were similar to those of laypeople, except that designers had a greater dislike for images consisting of lawns and retention ponds compared to laypeople, and designers had a slightly greater preference for newer forms of GSI, such as green roofs and bio-retentions. Designers can use these results to create preferred landscapes that manage stormwater while promoting the wellbeing of the people they serve.

Translation 1:

لقد نفذت العديد من المدن في الولايات المتحدة البنية التحتية الخضراء لمياه الأمطار لإدارة مياه الأمطار. ونظرًا لأن العديد من أنواع البنية التحتية الخضراء لمياه الأمطار جديدة نسبيًا، فإننا لا نعرف كيف ينظر الناس إلى البنية التحتية الخضراء لمياه الأمطار أو إلى أي مدى يفضلون تكوينات مختلفة من البنية التحتية الخضراء لمياه الأمطار. كما أننا لا نعرف إلى أي مدى ينظر عامة الناس إلى

البنية التحتية الخضراء لمياه الأمطار بشكل مختلف عن المصممين. ويكلفنا هذا الافتقار إلى المعرفة: فنحن نخاطر بإنشاء البنية التحتية الخضراء لمياه الأمطار التي يجدها الناس غير جذابة، وإذا لم يعجبهم ذلك، فقد لا يرغبون في قضاء الوقت فيها أو بالقرب منها، وقد لا يحافظون عليها، وقد لا يحصدون فوائد صحية منها. في هذه الدراسة، قدم 497 شخصًا عاديًا و117 مصممًا تقييمات تفضيلية لـ 55 صورة للبنية التحتية الخضراء لمياه الأمطار. لقد استخدمنا تحليل العوامل لتحديد فئات البنية التحتية الخضراء لمياه الأمطار التي يدركها الناس. لقد وجدنا أن المناظر الطبيعية التي تحتوي على أي بنية تحتية خضراء لمياه الأمطار مفضلة بشكل كبير من المناظر الطبيعية التي لا تحتوي على بنية تحتية خضراء لمياه الأمطار، وأن فوضى البنية التحتية الخضراء لمياه الأمطار أثرت على التفضيل. كانت تفضيلات المصممين مماثلة لتفضيلات عامة الناس، باستثناء أن المصممين لديهم كره أكبر للصور التي تتكون من حدائق وبرك احتجاز مقارنة بعامة الناس، وكان لدى المصممين تفضيل أكبر قليلاً للأشكال الأحدث من GSI، مثل الأسطح الخضراء والاحتجازات البيولوجية. يمكن للمصممين استخدام هذه النتائج لإنشاء مناظر طبيعية مفضلة تدير مياه الأمطار مع تعزيز رفاهية الأشخاص الذين يخدمونهم..

Translation 2:

قامت العديد من المدن في الولايات المتحدة بتطبيق بنية تحتية خضراء لمياه العواصف (GSI) لإدارة مياه الأمطار. نظرًا لأن عدة أنواع من هذه البنية التحتية جديدة نسبيًا، لا نعرف كيف ينظر الناس إليها أو إلى أي مدى يفضلون التكوينات المختلفة لها. كما أننا لا نعرف إلى أي مدى تختلف تصورات عامة الناس عن هذه البنية التحتية مقارنة بتصورات المصممين. هذا النقص في المعرفة يكلفنا: هناك خطر في إنشاء بنية تحتية قد يجدها الناس غير جذابة، وإذا لم تعجبهم، فقد لا يرغبون في قضاء الوقت فيها أو بالقرب منها، وقد لا يقومون بصيانتها، وبالتالي قد لا يحصلون على الفوائد الصحية منها. في هذه الدراسة، قدم 497 شخصًا من عامة الناس و117 مصممًا تقييمات تفضيلية لـ 55 صورة للبنية التحتية الخضراء. استخدمنا تحليل العوامل لتحديد الفئات التي يدركها الناس ضمن هذه البنية التحتية. ووجدنا أن المناظر الطبيعية التي تحتوي على أي بنية تحتية خضراء أكثر تفضيلاً بشكل ملحوظ من تلك التي لا تحتوي على أي بنية تحتية خضراء، وأن مدى الفوضى في التصميم أثر على التفضيل. كانت تفضيلات المصممين مشابهة لتفضيلات عامة الناس، باستثناء أن المصممين أظهروا نفورًا أكبر من الصور التي تحتوي على المروج وبرك الاحتجاز مقارنة بعامة الناس، وكان لديهم تفضيل طفيف للأشكال الأحدث من البنية التحتية الخضراء مثل الأسطح الخضراء والاحتفاظ البيولوجي. يمكن للمصممين استخدام هذه النتائج لإنشاء مناظر طبيعية مفضلة تُدير مياه الأمطار وتُعزز رفاهية الناس الذين يخدمونهم.

Translation 3:

قامت العديد من المدن في الولايات المتحدة بتنفيذ بنية تحتية خضراء لإدارة مياه الأمطار .(GSI) نظرًا لأن عدة أنواع من هذه البنية التحتية جديدة نسبيًا، لا نعرف كيف يدرك الناس هذه البنية أو إلى أي مدى يفضلون التكوينات المختلفة لها. كما أننا لا نعرف إلى أي مدى يدرك الأشخاص العاديون هذه البنية بشكل مختلف عن المصممين. هذا النقص في المعرفة يسبب لنا بعض الخسائر: فقد نخاطر بإنشاء بنية تحتية خضراء يجدها الناس غير جذابة، وإذا لم تعجبهم، قد لا يرغبون في قضاء الوقت فيها أو بالقرب منها، وقد لا يقومون بصيانتها، وبالتالي قد لا يحصلون على الفوائد الصحية التي توفرها. في هذه الدراسة، قدم 497 شخصًا من العامة و117 مصممًا تقييمات تفضيلية لـ 55 صورة للبنية التحتية الخضراء. استخدمنا تحليل العوامل لتحديد الفئات التي يدركها الناس في هذه البنية. وجدنا أن المناظر الطبيعية التي تحتوي على أي نوع من البنية التحتية الخضراء مفضلة بشكل ملحوظ عن تلك التي لا تحتوي على أي منها، وأن الفوضى في تصميم البنية التحتية أثرت على التفضيلات. كانت تفضيلات المصممين مشابهة لتفضيلات العامة، باستثناء أن المصممين كانوا أقل تفضيلًا للصور التي تحتوي على المروج وبرك الاحتجاز مقارنة بالعامة، كما أن المصممين أبدوا تفضيلًا أكبر قليلًا للأشكال الأحدث من البنية التحتية الخضراء، مثل الأسطح الخضراء والأحواض البيولوجية. يمكن للمصممين استخدام هذه النتائج لإنشاء مناظر طبيعية مفضلة تدير مياه الأمطار وتدعم رفاهية الأشخاص الذين يخدمونهم.

Translation 4:

أقامت العديد من المدن الأمريكية أنظمة خضراء لتصريف مياه الأمطار. ولأن أنواعًا كثيرة من هذه الأنظمة حديثة نسبياً، فإننا لا نعرف كيف ينظر الناس لها أو مدى تفضيلهم لأشكال أخرى من هذه الأنظمة، ولا نعرف حتى مدى اختلاف تصورات العامة عن هذه الأنظمة مقارنة بمصممي هذه الأنظمة. وبالطبع، لهذا الجهل خسائر؛ إذ نخاطر بإنشاء بنية تحتية قد لا ينجذب الناس لها مما يعني أنهم قد لا يجدون في أنفسهم الرغبة في تمضية الوقت فيها أو بجوارها أو الحفاظ عليها وربما لا يحصلون على الفوائد الصحية المرجوة منها. وفي هذه الدراسة، قَيَّم 497 شخصًا من غير المتخصصين و 117 مصممًا متخصصًا 55 صورة من صور الأنظمة الخضراء لتصريف مياه الأمطار. ولكي نحدد فئات الأنظمة التي يستطيع العوام التعرف عليها، استخدمنا تحليلًا إحصائيًا ألا وهو التحليل العاملي الذي قادنا إلى نتائج عدة، أهمها أن العوام يفضلون المناظر الطبيعية المزودة بأي من أنواع الأنظمة الخضراء لتصريف مياه الأمطار أكثر بكثير من تلك التي لا تحتوي على أي من هذه الأنظمة وأن العشوائية في بنية هذه الأنظمة كان لها تأثير بارز على قرارارت هؤلاء العوام. أما آراء المصممين فقد جاءت مشابهة لآراء العوام وإن خالفوهم في نقطتين، حيث أبدى المصممون نفورًا أكبر تجاه الصور التي تظهر بها المروج والبرك الصناعية، في حين زاد إعجابهم قليلًا بالأشكال الأكثر حداثة من الأنظمة الخضراء لتصريف مياه الأمطار مثل الأسطح الخضراء وحدائق الأمطار. وبهذا، يمكن للمصممين أن يستندوا لهذه النتائج في تصميم المناظر الطبيعية التي يفضلها العوام لتساهم في إدارة مياه الأمطار وتحسن من رفاهيتهم.

==========

Section 5: Evaluation of Text 4 Translations: Please review the following text and its four Arabic translations. Then, evaluate each based on the questions provided below.

Source Text:

Showdown is set: Trump and Harris plan to debate in September

ABC News has said it will host the first debate between Donald Trump and Kamala Harris on 10 September.

The network confirmed the showdown in a post on X, formerly Twitter, after Trump said on Thursday that he would be open to debating his Democratic rival multiple times before the November election.

"We think we should do three debates," Trump said, suggesting two additional debates that he said would be hosted by Fox News and NBC, respectively.

Ms. Harris confirmed that she will attend the ABC debate while at an event in Michigan on Thursday, and said later that she would be open to additional debates.

The network said the debate will be moderated by World News Tonight anchor and managing editor David Muir and ABC News Live Prime anchor Linsey Davis.

Translation 1:

المواجهة محددة: ترامب وهاريس يخططان للمناظرة في سبتمبر

قالت شبكة ABC News إنها ستستضيف المناظرة الأولى بين دونالد ترامب وكامالا هاريس في 10 سبتمبر.

وأكدت الشبكة المواجهة في منشور على X، المعروف سابقًا باسم Twitter، بعد أن قال ترامب يوم الخميس إنه سيكون منفتحًا على مناظرة منافسه الديمقراطي عدة مرات قبل انتخابات نوفمبر.

وقال ترامب: "نعتقد أنه يجب علينا إجراء ثلاث مناظرات"، مقترحًا مناظرتين إضافيتين قال إنهما ستستضيفهما Fox News وNBC على التوالي.

وأكدت السيدة هاريس أنها ستحضر مناظرة ABC أثناء وجودها في حدث في ميشيغان يوم الخميس، وقالت لاحقًا إنها ستكون منفتحة على مناقشات إضافية.

وقالت الشبكة إن المناظرة سيديرها مذيع World News Tonight ورئيس التحرير ديفيد موير ومذيعة ABC News Live Prime لينسي ديفيس.

Translation 2:

المواجهة قد تحددت: ترامب وهاريس يخططان لمناظرة في سبتمبر

أعلنت شبكة ABC News أنها ستستضيف أول مناظرة بين دونالد ترامب وكمالا هاريس في 10 سبتمبر.

وأكدت الشبكة المواجهة في منشور على منصة X (المعروفة سابقًا باسم تويتر)، بعد أن قال ترامب يوم الخميس إنه مستعد لمناظرات متعددة مع منافسته الديمقراطية قبل انتخابات نوفمبر.

وقال ترامب: "نعتقد أنه يجب أن تكون هناك ثلاث مناظرات"، مقترحًا تنظيم مناظرتين إضافيتين تستضيفهما قناة فوكس نيوز و NBCعلى التوالي.

وأكدت السيدة هاريس حضورها مناظرة ABC خلال حدث في ميشيغان يوم الخميس، وأعلنت لاحقًا أنها ستكون مستعدة لمزيد من المناظرات.

وذكرت الشبكة أن المناظرة ستديرها مذيعة برنامج "وورلد نيوز توناي" ورئيسة التحرير ديفيد موير، والمذيعة لينسي ديفيس من ABC News Live Prime .

Translation 3:

المواجهة محددة: ترامب وهاريس يخططان لمناظرة في سبتمبر

أعلنت قناة ABC News أنها ستستضيف أول مناظرة بين دونالد ترامب وكامالا هاريس في 10 سبتمبر. وأكدت الشبكة المواجهة في منشور على منصة X (تويتر سابقًا) بعد أن صرح ترامب يوم الخميس بأنه سيكون منفتحًا على مناظرات متعددة مع منافسته الديمقراطية قبل انتخابات نوفمبر.

وقال ترامب: "نعتقد أنه يجب أن تكون هناك ثلاث مناظرات"، مقترحًا مناظرتين إضافيتين يستضيفهما على التوالي Fox News و.NBC

وأكدت السيدة هاريس حضورها في مناظرة ABC أثناء مشاركتها في حدث في ميشيغان يوم الخميس، وذكرت لاحقًا أنها ستكون منفتحة على مناظرات إضافية.

وذكرت الشبكة أن المناظرة سيشرف عليها مذيع برنامج "World News Tonight" ورئيس تحريرABC News ، ديفيد موير، بالإضافة إلى مذيعة "ABC News Live Prime" لينسي ديفيس.

Translation 4:

مناظرة بين ترامب وهاريس في سبتمبر المقبل

قالت شبكة "إيه بي سي نيوز" الأمريكية إنها ستستضيف المناظرة الأولى بين المرشحين الرئاسيين دونالد ترامب وكامالا هاريس في العاشر من سبتمبر/أيلول.

وأكدت الشبكة هذه المواجهة في منشور على إكس، المعروف سابقاً باسم تويتر، بعد أن قال ترامب يوم الخميس إنه سيكون منفتحاً على مناظرة منافسته الديمقراطية عدة مرات قبل انتخابات نوفمبر/تشرين الثاني.

وقال ترامب: "نعتقد أنه ينبغي علينا إجراء ثلاث مناظرات"، مقترحاً مناظرتين إضافيتين على أن تستضيفهما فوكس نيوز وإن بي سي على التوالي.

وأكدت هاريس أنها ستعقد مناظرة "إيه بي سي" أثناء وجودها في ميشيغان يوم الخميس، وقالت لاحقاً إنها ستكون منفتحة على المزيد من المناظرات.

وقالت الشبكة إن المناظرة سيديرها ديفيد موير، المذيع ومدير تحرير البرنامج الإخباري "وورلد نيوز توناية" أو "أخبار العالم الليلة"، ولينسي ديفيس، المذيعة في البرنامج التحليلي "إيه بي سي نيوز لايف برايم".

1. How far does each translation accurately convey the meaning of the source text? Please respond to the question by marking the box you find appropriate with an "X."

		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
Text 1	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
Text 2	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
Text 3	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
Text 4	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					

2. How far is each translation easy to understand and follow? Please respond to the question by marking the box you find appropriate with an "X."

		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
Text 1	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					
Text 2		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1

	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
	Translation 1					
Text 3	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
	Translation 1					
Text 4	Translation 2					
	Translation 3					
	Translation 4					

3. How far does each translation match the style and tone of the source text? Please respond to the question by marking the box you find appropriate with an "X."

		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
	Translation 1					
Text 1	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
	Translation 1					
Text 2	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
	Translation 1					
Text 3	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
	Translation 1					
Text 4	Translation 2					
	Translation 3					
	Translation 4					

4. How far is each translation culturally appropriate for the target audience? Please respond to the question by marking the box you find appropriate with an "X."

		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
	Translation 1					
Text 1	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
	Translation 1					
Text 2	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
	Translation 1					
Text 3	Translation 2					
	Translation 3					
	Translation 4					
Text 4		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1

	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					

5. How far does each translation use specialized terminology accurately and consistently? Please respond to the question by marking the box you find appropriate with an "X."

		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
Text 1	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
Text 2	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
Text 3	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					
		Completely 5	Mostly 4	Somewhat 3	Slightly 2	Not at all 1
Text 4	Translation 1					
	Translation 2					
	Translation 3					
	Translation 4					

Kindly include any additional comments you may have:

_____
_____
_____
_____
_____
_____

Appendix 2: Human evaluators' ratings of the translations under study

Evaluator	Text Type	Translation done via	Accuracy	Fluency	Style & Tone	Cultural suitability	Terminology	Averages
Evaluator1	General	Google Translate	3	1	1	2	2	1.8
Evaluator2			4	5	4	3	5	4.2
Evaluator3			3	3	2	2	2	2.4
Evaluator4			4	5	4	4	4	4.2
Evaluator5			4	4	3	4	4	3.8
Evaluator6			2	2	2	3	2	2.2
Evaluator7			4	4	4	4	4	4
Evaluator8			3	3	3	4	3	3.2
Evaluator9			3	3	3	3	3	3
Evaluator10			2	3	2	3	2	2.4
	Averages		**3.2**	**3.3**	**2.8**	**3.2**	**3.1**	**3.12**
Evaluator1	General	ChatGPT simple-prompt	2	2	2	3	2	2.2
Evaluator2			5	5	5	4	5	4.8
Evaluator3			4	3	2	2	2	2.6
Evaluator4			4	5	4	4	4	4.2
Evaluator5			2	4	3	3	4	3.2
Evaluator6			3	3	3	2	3	2.8
Evaluator7			4	4	4	4	4	4
Evaluator8			4	3	4	3	3	3.4
Evaluator9			4	4	4	4	4	4
Evaluator10			3	3	3	3	2	2.8
	Averages		**3.5**	**3.6**	**3.4**	**3.2**	**3.3**	**3.4**
Evaluator1	General	ChatGPT detailed-prompt	4	3	4	4	3	3.6
Evaluator2			3	5	3	2	5	3.6
Evaluator3			4	4	3	4	4	3.8
Evaluator4			4	5	4	4	4	4.2
Evaluator5			4	4	4	4	4	4
Evaluator6			4	4	4	4	5	4.2
Evaluator7			4	4	4	4	4	4
Evaluator8			5	3	3	3	3	3.4
Evaluator9			4	4	4	5	4	4.2
Evaluator10			4	4	4	4	4	4
	Averages		**4**	**4**	**3.7**	**3.8**	**4**	**3.9**
Evaluator1	General	Human translation	5	4	3	5	5	4.4
Evaluator2			2	5	2	5	5	3.8
Evaluator3			5	5	4	5	4	4.6
Evaluator4			5	5	5	5	5	5
Evaluator5			5	4	5	4	4	4.4
Evaluator6			5	5	5	5	4	4.8
Evaluator7			5	5	5	5	5	5
Evaluator8			3	3	3	4	3	3.2
Evaluator9			5	5	4	5	5	4.8
Evaluator10			4	5	4	5	5	4.6
Evaluator1	**Averages**		**4.4**	**4.6**	**4**	**4.8**	**4.5**	**4.46**
Evaluator1	Literary	Google Translate	2	2	2	1	1	1.6
Evaluator2			5	5	5	4	5	4.8
Evaluator3			4	4	3	3	3	3.4
Evaluator4			4	5	4	4	4	4.2

Evaluator5			4	4	5	4	4	4.2
Evaluator6			3	3	3	3	3	3
Evaluator7			5	3	5	3	3	3.8
Evaluator8			4	4	4	4	4	4
Evaluator9			5	5	5	3	4	4.4
Evaluator10			2	4	3	4	3	3.2
	Averages		**3.8**	**3.9**	**3.9**	**3.3**	**3.4**	**3.66**
Evaluator1	Literary	ChatGPT simple-prompt	3	2	3	2	2	2.4
Evaluator2			4	3	4	2	3	3.2
Evaluator3			4	5	4	4	4	4.2
Evaluator4			4	5	4	4	4	4.2
Evaluator5			4	4	5	4	3	4
Evaluator6			3	4	3	3	3	3.2
Evaluator7			5	3	5	3	3	3.8
Evaluator8			3	4	4	4	4	3.8
Evaluator9			5	4	4	4	4	4.2
Evaluator10			4	4	4	4	4	4
	Averages		**3.9**	**3.8**	**4**	**3.4**	**3.4**	**3.7**
Evaluator1	Literary	ChatGPT detailed-prompt	3	3	3	3	3	3
Evaluator2			3	4	4	5	4	4
Evaluator3			4	5	4	4	4	4.2
Evaluator4			4	5	4	4	4	4.2
Evaluator5			4	4	5	4	3	4
Evaluator6			4	5	4	4	4	4.2
Evaluator7			5	4	4	4	4	4.2
Evaluator8			4	4	5	4	4	4.2
Evaluator9			5	4	4	4	4	4.2
Evaluator10			5	5	5	4	5	4.8
	Averages		**4.1**	**4.3**	**4.2**	**4**	**3.9**	**4.1**
Evaluator1	Literary	Human translation	5	5	5	5	5	5
Evaluator2			2	2	4	3	2	2.6
Evaluator3			4	4	3	3	3	3.4
Evaluator4			5	5	5	5	5	5
Evaluator5			4	5	5	4	4	4.4
Evaluator6			5	5	5	5	5	5
Evaluator7			5	5	3	5	5	4.6
Evaluator8			4	4	4	4	4	4
Evaluator9			5	5	3	5	5	4.6
Evaluator10			4	5	4	5	4	4.4
	Averages		**4.3**	**4.5**	**4.1**	**4.4**	**4.2**	**4.3**
Evaluator1	Scientific	Google Translate	2	2	2	2	2	2
Evaluator2			5	2	5	2	5	3.8
Evaluator3			1	1	1	1	1	1
Evaluator4			3	3	3	4	4	3.4
Evaluator5			3	3	4	3	3	3.2
Evaluator6			3	3	3	3	3	3

Evaluator7			3	2	5	2	2	2.8
Evaluator8			4	4	3	4	3	3.6
Evaluator9			4	4	4	3	4	3.8
Evaluator10			2	2	2	3	2	2.2
	Averages		**3**	**2.6**	**3.2**	**2.7**	**2.9**	**2.88**
Evaluator1	Scientific	ChatGPT simple-prompt	3	3	3	3	3	3
Evaluator2			4	3	4	3	4	3.6
Evaluator3			2	2	2	2	2	2
Evaluator4			4	5	3	4	4	4
Evaluator5			3	3	4	3	3	3.2
Evaluator6			2	2	3	3	3	2.6
Evaluator7			3	3	4	3	3	3.2
Evaluator8			5	4	3	4	3	3.8
Evaluator9			4	4	4	3	4	3.8
Evaluator10			4	3	4	3	2	3.2
	Averages		**3.4**	**3.2**	**3.4**	**3.1**	**3.1**	**3.24**
Evaluator1	Scientific	ChatGPT detailed-prompt	4	4	5	4	4	4.2
Evaluator2			3	5	3	5	3	3.8
Evaluator3			4	4	4	4	4	4
Evaluator4			4	5	4	4	4	4.2
Evaluator5			4	4	3	4	4	3.8
Evaluator6			4	4	4	4	4	4
Evaluator7			4	4	3	4	3	3.6
Evaluator8			5	4	3	3	3	3.6
Evaluator9			5	5	4	4	4	4.4
Evaluator10			4	4	4	4	3	3.8
Evaluator1	**Averages**		**4.1**	**4.3**	**3.7**	**4**	**3.6**	**3.94**
Evaluator1	Scientific	Human translation	5	5	4	5	5	4.8
Evaluator2			2	4	2	4	2	2.8
Evaluator3			4	4	4	4	4	4
Evaluator4			5	5	5	5	5	5
Evaluator5			4	4	4	4	4	4
Evaluator6			5	5	5	5	5	5
Evaluator7			5	5	2	5	5	4.4
Evaluator8			4	4	3	3	3	3.4
Evaluator9			5	5	4	5	4	4.6
Evaluator10			4	5	4	5	4	4.4
	Averages		**4.3**	**4.6**	**3.7**	**4.5**	**4.1**	**4.24**
Evaluator1	Media	Google Translate	4	3	3	2	2	2.8
Evaluator2			5	4	5	2	5	4.2
Evaluator3			3	3	1	2	1	2
Evaluator4			5	5	5	5	4	4.8
Evaluator5			5	5	4	4	5	4.6
Evaluator6			3	3	3	3	3	3
Evaluator7			4	4	5	3	3	3.8
Evaluator8			4	5	4	5	4	4.4
Evaluator9			5	5	4	3	4	4.2
Evaluator10			2	2	3	2	2	2.2
	Averages		**4**	**3.9**	**3.7**	**3.1**	**3.3**	**3.6**
Evaluator1	Media	ChatGPT simple-prompt	4	4	4	3	4	3.8
Evaluator2			4	4	4	3	4	3.8
Evaluator3			4	4	2	3	3	3.2
Evaluator4			5	3	5	5	4	4.4
Evaluator5			5	5	4	4	4	4.4
Evaluator6			5	5	5	5	5	5
Evaluator7			4	4	4	4	4	4
Evaluator8			4	5	4	5	4	4.4
Evaluator9			5	5	4	4	4	4.4
Evaluator10			2	3	3	3	2	2.6
	Averages		**4.2**	**4.2**	**3.9**	**3.9**	**3.8**	**4**

Evaluator1	Media	ChatGPT detailed-prompt	4	4	4	4	4	4
Evaluator2			4	4	3	4	4	3.8
Evaluator3			5	5	4	4	4	4.4
Evaluator4			5	5	5	5	4	4.8
Evaluator5			5	5	4	4	5	4.6
Evaluator6			3	3	3	3	3	3
Evaluator7			5	4	4	4	4	4.2
Evaluator8			4	5	4	4	4	4.2
Evaluator9			5	5	3	4	4	4.2
Evaluator10			3	3	4	2	3	3
	Averages		**4.3**	**4.3**	**3.8**	**3.8**	**3.9**	**4.02**
Evaluator1	Media	Human translation	5	5	4	5	5	4.8
Evaluator2			4	4	2	5	3	3.6
Evaluator3			5	5	4	4	4	4.4
Evaluator4			5	5	5	5	5	5
Evaluator5			5	5	4	4	4	4.4
Evaluator6			4	4	4	4	4	4
Evaluator7			5	5	3	5	5	4.6
Evaluator8			4	5	4	5	4	4.4
Evaluator9			5	5	3	5	5	4.6
Evaluator10			4	5	3	5	5	4.4
	Averages		**4.6**	**4.8**	**3.6**	**4.7**	**4.4**	**4.42**