A Translation Quality Evaluation Dimensions

As presented in Table 1, we adopt 11 core MQM error categories[1] to evaluate game localization translation quality.

B Offensive Language Evaluation Dimensions

We identify 430 offensive language cases, categorized into different types and factors according to Table 2.

C Evaluation Guideline

Thank you for participating in the evaluation task. The goal of this annotation task is to assess the quality of translations and to evaluate how offensive language is conveyed during translation in the game Black Myth: Wukong.

General Instructions: You will be provided with a source text in Chinese and its English translation. Your task consists of two main components. Please annotate carefully and consistently, adhering to the criteria outlined below to ensure fairness and objectivity:

- *Evaluate Translation Quality: Assess the translation based on three quality dimensions: accuracy, fluency, and audience appropriateness, using the specified criteria and a 5-point Likert scale.*

- *Rate Language Offensiveness: Determine whether the offensive content in the translation has been intensified, softened, or remains neutral compared to the source text.*

Task 1: Translation Quality Evaluation For this task, you will assess the quality of a given translation by comparing it to the original Chinese text. The evaluation involves three key dimensions: accuracy, fluency, and audience appropriateness. Each dimension will be rated on a 5-point Likert scale, with specific criteria for scoring.

- *1 = Poor: The translation is unusable. It fails to convey the meaning of the source text, contains major errors, or is entirely incoherent.*

- *2 = Fair: The translation contains significant issues affecting its accuracy, fluency, or appropriateness. While some meaning may be conveyed, it is highly flawed and difficult to use.*

- *3 = Average: The translation conveys the general meaning of the source text but contains noticeable issues. Usable with revisions.*

- *4 = Good: The translation is accurate, fluent, and appropriate for the audience, with only minor issues that do not significantly affect comprehension or usability.*

- *5 = Excellent: The translation is near-perfect, fully accurate, fluent, and appropriate for the audience, with no noticeable errors.*

Accuracy: Evaluate how well the translation conveys the meaning of the source text. Check for errors such as mistranslations, omissions, or unnecessary additions that alter the original meaning. A translation should accurately convey both the content and intent of the source text.

Fluency: Assess the grammatical correctness, naturalness, and stylistic quality of the translation in English.

Audience Appropriateness: Evaluate whether the translation is suitable for the intended audience, in particular, the culture-specific terms.

See Table 1 for detailed error category.

Task 2: Offensiveness Degree Rating

In this task, you will evaluate whether the translation amplifies, softens, or faithfully retains the offensive content of the source text. Each translation should be classified based on the level of offensiveness relative to the source text.

More Offensive: The translation is more offensive than the source text. This includes cases with amplified offensive intensity, stronger offensive connotations, and escalated tone.

Neutral: The translation maintains the same level of offensiveness as the source text. This applies when the emotional intensity remains consistent, and the translation faithfully preserves the tone and intent of the original.

Less Offensive: The translation is less offensive than the source text. This includes cases with reduced offensive intensity, milder offensive connotations, and softened tone.

---

[1] https://themqm.org/error-types-2/typology/

Type	Definition	Error Category	Explanation
Accuracy	Errors occurring when the target content does not accurately correspond to the propositional content of the source text because of distortion, omission, or addition to the message.	Mistranslation	Incorrect translation that deviates from the source text.
		Addition	Added information not present in the source text.
		Omission	Missing information from the source text.
		Untranslated	Not translated.
		Repetition	Repeated or redundant translation.
Fluency	Errors related to the form or content of a text, irrespective of whether it is a translation or not.	Grammar	Grammatical errors.
		Punctuation	Incorrect or missing punctuation.
		Spelling	Misspelled words or typographical errors.
		Language Register	Inappropriate tone or style.
		Flow	Lack of natural flow or logical connections.
Audience Appropriateness	Errors arising from the use of target content that is invalid or inappropriate for the target locale or target audience.	Culture-specific Reference	Mistranslated or inappropriate cultural terms.

Table 1: Adapted MQM framework with 11 core categories, balancing comprehensiveness with annotation efficiency.

Dimension	Explanation	Example
Derogatory Insult	Negative evaluations attacking appearance, behavior, abilities, intelligence, or character.	[ZH] 你这般不识抬举，算不算背恩忘情 [EN] Yet you spurned each one. Such ingratitude!
Sarcasm and Irony	Opposite meaning to literal interpretation, often in a mocking or critical tone.	[ZH] 有功的猴子 [EN] A monkey of merit
Tone and Context	Emotional undertone alters neutral words.	[ZH] 小外甥，这就是你说的单挑？ [EN] Brat, you call that a duel?
Gender Bias	Gender biases and stereotypes.	[ZH] 少不得要再把那小贱人抓回来，好好耍子一番 [EN] That little fox, I'll hunt her down
Religious Demeaning	Demeans religious beliefs.	[ZH] 信甚么狗屁如来，不如我自己来 [EN] Why bet on the Sutra, when one oneself can be a Buddha?
Violation of Social Morals	Violations of social norms and structures, including familial deference and institutional authority.	[ZH] 莫急啊，老畜生 [EN] Worry not, old, horned fool.
Offensive Vernacular	Offensive colloquial expressions and slang violating social etiquette.	[ZH] 奶奶的，碍事！ [EN] Curse it. You're in my way!
Culture-specific Connotations	Terms or idioms carrying specific offensive connotations in *Journey to the West* and Chinese culture.	[ZH] 弼马温，你不是死了么！ [EN] Filthy monkey! Aren't you dead?

Table 2: Examples for identifying different types of offensive language.