

Analyzing Translations of Abstracts in Physical Education Theses Based on the MQM: Human versus AI-Generated Translations¹

Narjes Asgari Vartooni², Mahmoud Afrouz³
and Vazgen Minasian⁴

Abstract

The study focuses on analyzing the quality of abstract translations in Physical Education theses. It aimed to identify errors in human versus AI-generated translations. The paper also aimed to explore human-generated translations using the Multidimensional Quality Metrics (MQM) framework. For this purpose, 30 English abstracts from the field of sports physiology, spanning 15 years (1390s and 1400s/ 2011 to 2025), were selected. Errors were categorized and quantified by type and frequency. The analysis revealed a fluctuating pattern, characterized by a significant recent improvement. Results showed a general upward trend in errors from 2011 to 2020, including a gradual decline in abstract quality—a trend potentially linked to an over-reliance on machine translation. This pattern reversed sharply in the 2020s, with data showing a consistent and marked reduction in errors, culminating in a 71% decrease by 1404 (2025) and reflecting a substantial enhancement in translation quality during the early years of the 1400s (2021–2025).

Keywords: Artificial intelligence; Multidimensional Quality Metrics; Physical education; Sports physiology; Translation quality assessment

1. This paper was received on 26.12.2025 and approved on 18.04.2026.

2. MA Student of Translation Studies, Department of English Language and Literature, University of Isfahan), Isfahan, Iran; email: na014asgari@fgn.ui.ac.ir

3. Corresponding Author: Associate Professor, Department of English Language and Literature University of Isfahan) , Isfahan, Iran; email: m.afrouz@fgn.ui.ac.ir

4. Professor, Department of Exercise Physiology, Faculty of Sports Sciences, University of Isfahan, Isfahan, Iran; email: v.minasian@spr.ui.ac.ir

Introduction

A relatively neglected domain in the assessment of scholarly texts pertains to the qualitative evaluation of translations of abstracts. Under Iran's system of higher education, students are required to provide an English abstract along with a Persian abstract in their thesis in order to present their research findings to the rest of the world.

Translation quality assessment (TQA), in general, has gained prominence in product-focused translation and attracted significant interest from both scholars and professional translators (Manipuspika, 2021). However, there has been limited research on evaluating the translation of academic texts, such as thesis and article abstracts (Jafarian, 2016; Mahdavi, 2017), while the existing literature on TQA in other genres is quite extensive and substantial (e.g., Munawir, 2020; Derakhshi & KhorramiFard, 2019; Napitupulu, 2017; Angelelli, 2009).

The examination of translation errors in student abstracts offers helpful information. These errors are often the result of persisting linguistic weaknesses, for example, technical terminology, grammar, or syntactical organization. Focusing on these areas can assist students in enhancing their language competence—so not only improving the level of their present work but also getting them ready for future academic publication opportunities (Asgari Vartooni et al., 2026). When theses abstracts are translated poorly and contain a multitude of linguistic or structural errors, even the academic reputation of the university could be damaged. Although translation should not be deemed as a “purely technical skill” and, therefore “clear-cut solutions” cannot be found for all translation problems (Asgari Vartooni, 2026: 2), detecting errors occurred in translated texts, categorizing them and discussing the possible reasons behind their occurrences may reduce translation problems.

The advent of artificial intelligence technologies has led to increased reliance on these tools among students. Nonetheless, it remains uncertain whether such technologies can consistently produce accurate and high-quality translations in specialized fields like Physical Education. The purpose of this study is to compare the quality of abstract translations of Physical Education master's theses over the last 15 years (decades 1390s–1400s). The study aims at answering the following questions:

1. What are the most frequent errors made in human-generated translations compared to AI generated translations?
2. How has the quality of human-generated abstract translations in Physical

Education theses changed over the past 15 years?

Literature Review

The notion of quality seems to be a sensitive and complex issue within translation studies. Error-based models for TQA focus on finding and classifying different types of errors in a translation, which can be linguistic, cultural or stylistic. Developed by Lommel et al (2014), MQM is a flexible framework for assessing translation quality. It is utilized extensively in academic contexts. What follows is a review of related previous studies.

Bam (2005) carried out an exploration that examines similarities and differences between sports texts originally written in Afrikaans and sports texts translated from English into Afrikaans. The investigation shows that these dissimilarities are not statistically significant. In Löfgren's (2014) study, the translation of three sports texts from English into Swedish is explored, with particular attention given to the translation procedures employed when translating sports terminology. The author simply emphasized the importance of using proper procedures while rendering the sports terminology.

Terzi and Arslantürk (2014) focused on evaluating English abstracts of MA and PhD dissertations published in Turkish and identifying translation errors. A random selection of abstracts from both Turkish-speaking and English-speaking researchers was used. The authors point out that the analysis of rhetorical moves did not reveal significant differences in move structures across the abstracts. The corpus, however, was too small to let them generalize their findings.

Boynukara (2017) has concentrated on the important role of translation and interpretation in international sports. The researcher has analyzed the ways in which these linguistic activities have contributed to making sports a global phenomenon. Boynukara (2017) points out that effective sports translation needs not only linguistic competence for sports translation but also a deep understanding of the sporting contexts. This is, however, true about other fields of knowledge and not something to be specifically concluded by this work.

Yuan (2019) explores the utilization of linguistic theories of translation in the context of sports English translation. Based on his findings, sports texts mainly involve six categories: sports performance, sports news, sports marketing, sports intelligence, sports education, and sports literature. Based on Yuan's (2019) purely theoretical

study, the use of linguistic theories of translation makes it feasible to define sports texts and their types and functions.

Li and Chen's (2019) study evaluated four prominent online MT platforms, focusing on their efficiency, operational modes, and conditions. The results indicate that, while MT quality is steadily improving, a significant gap remains between machine and human translations. Based on these findings, the author foresees that automated translation is unlikely to replace human translation in the near future.

Herdawan's (2020) study aimed to explore the translation strategies employed by Google Translate and the types of errors it made when translating Indonesian-English abstracts. A descriptive qualitative methodology was used by the researcher who analyzed 15 abstracts. The findings revealed that Google Translate primarily uses a literal translation strategy. It also made several linguistic errors and numerous translation errors.

Zandi (2020) investigates translational errors in article and thesis abstracts written by Iranian PhD students in psychology based on Waddington's model. The findings revealed that while omission and addition were the most common errors in both thesis and article abstracts, errors related to lexical items and inappropriate linguistic variation were the least frequent. A more contemporary TQA model could have been used by the researcher (e.g., MQM, ATA, etc.).

3. Methodology

The study employs a descriptive and corpus-based research design. The primary goal was to evaluate the evolution of translation quality in master's thesis abstracts within the field of physical education (PE) over the past 15 years. To ensure a representative sample, 15 abstracts were selected from each decade, resulting in a total of 30 abstracts.

At first abstracts were scanned and were then processed by OCR software to convert them into editable text. The Persian abstracts were consequently input into DeepSeek-V3.2 (free version, June 2025) for translation. The process was carried out employing a single prompt in a single session within a 24-hour period. In the subsequent phase, both human and AI-generated (Persian-to-English) translations were analyzed for error analysis based on MQM. To ensure the reliability of the classification process, two raters—both MA students in translation studies and familiar with MQM—were invited to review the data.

The MQM is a model for analyzing and evaluating translation quality. The last

version '2024_03-07-MQMFull_Master-Official' is used in this study (retrieved 17 October 2025 from <https://themqm.org/>). The top-level error dimensions of the MQM used for data analysis included Terminology, Accuracy, Linguistic conventions, Style, Locale conventions, Audience appropriateness, and Design and markup.

4. Results and Discussion

The section provides an analysis of the research findings. It starts with the results section which displays the data through descriptive tables and figures to highlight key patterns. The discussion section offers an interpretation of the relevant examples extracted from the corpus.

4.1. Results

To ensure a systematic presentation of data, the findings related to each decade are represented in separate tables.

4.1.1 Classification of errors in the 1390s

In table 4.1 frequency of the top level error dimensions are displayed in abstracts translated by human translator (HT) in 1390s (2011 – 2020). Blank cells in tables indicate zero frequency.

Table 4.1 Frequency of top-level error dimensions occurred in each year (in HT)

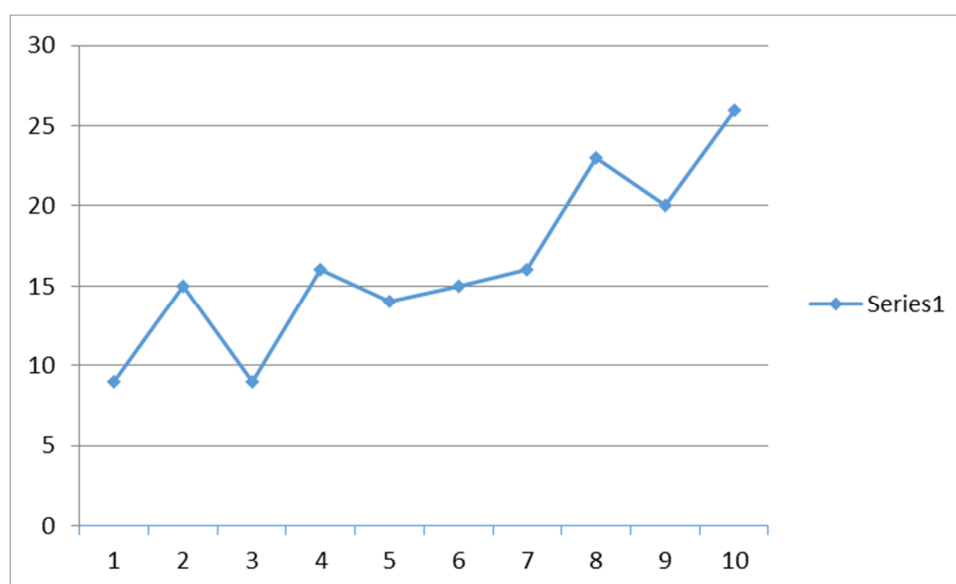
Years Error Dimensions	90	91	92	93	94	95	96	97	98	99	Total
Terminology	1		1	3	2	1	2	1		6	17
Accuracy	3	8	1	2	1	3	5	8	7	2	40
Linguistic conventions	5	6	6	7	10	11	7	13	13	16	94
Style		1	1	4	1		2	1		2	12
Locale conventions											
Audience appropriateness											
Design and markup											

The total frequency of top-level error dimensions occurred in each year (from 1390 to 1399/ 2011 to 2020) is displayed in Table 4.2.

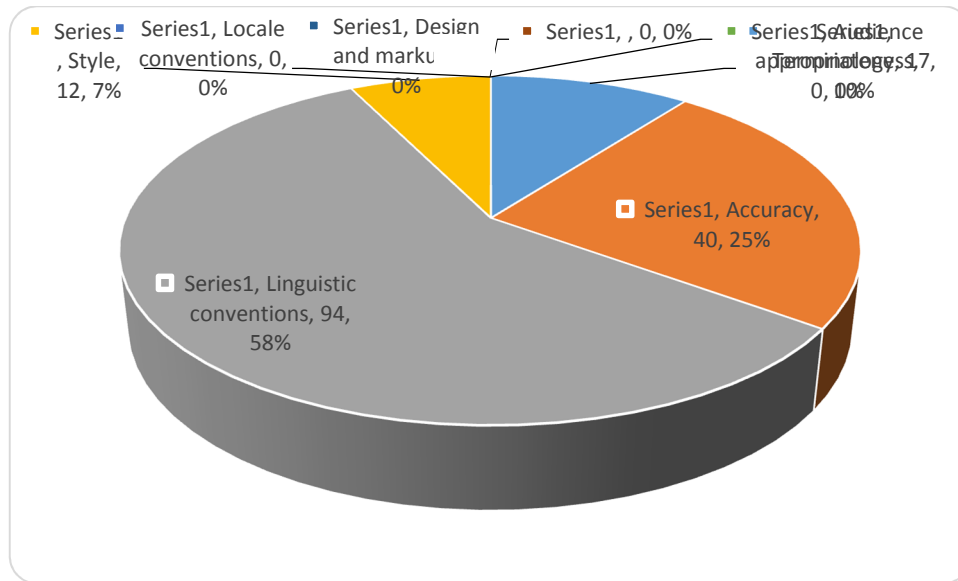
Table 4.2 *Total frequency of top-level error dimensions occurred in each year (in HT)*

Years	1	2	3	4	5	6	7	8	9	10
	(90)	(91)	(92)	(93)	(94)	(95)	(96)	(97)	(98)	(99)
Frequency	9	15	9	16	14	15	16	23	20	26
Percentage	5%	9%	6%	10%	9%	9%	10%	14%	12%	16%

The following line chart (Figure 4.1) is presented to make the general trend of error occurrence of in each year more tangible.

Figure 4.1 *Trend of top-level error dimensions occurred in 1390s (2011–2020)*

The data indicates a generally upward trend in the number of errors in English abstracts over the years from 1390 to 1399 (2011–2020). Although occasional declines are observed in certain years, such as 1392 and 1398, the overall pattern points to a gradual increase in error frequency. This rise may reflect growing challenges in maintaining abstract quality, possibly due to the over reliance of students on MTs. The total percentages of errors are displayed in Figure 4.2.

Figure 4.2 Top-level error dimensions occurred in 90s (2011–2020)

As is displayed in Figure 4.2, Terminology includes 10% of the errors. Accuracy also accounts for 25%. Linguistic conventions and Style include 58% and 7%, respectively. The remaining three categories register 0%. Only one Accuracy-related error is found in DeepSeek's translation.

4.1.2 Classification of errors in the 1400s

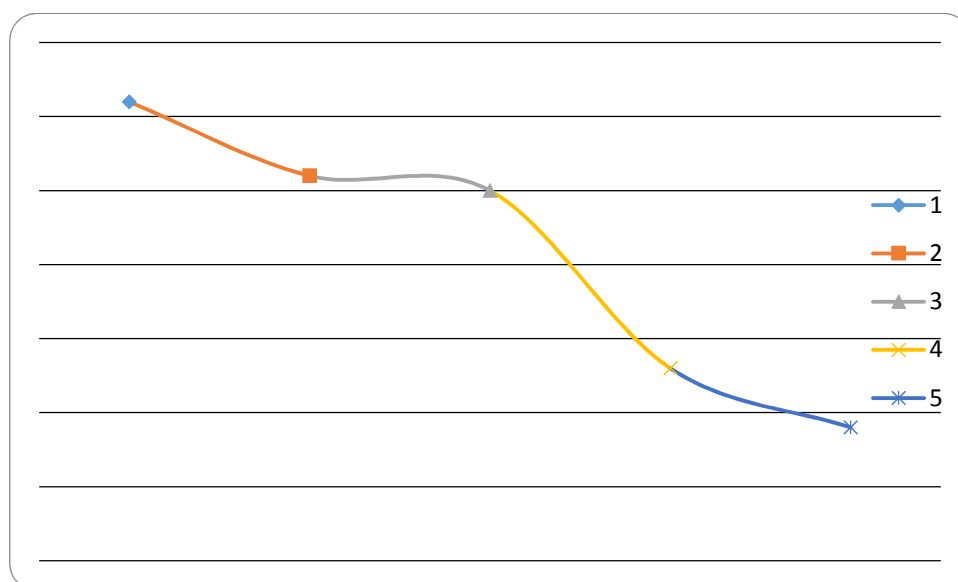
In the current section, only the human translation errors were analyzed due to the absence of errors in AI-generated translation. In fact, based on the MQM criteria and its top-level error dimensions, the researchers found no error in the AI-generated translation of abstract theses of 1400s.

Table 4.3 Frequency of top-level error dimensions occurred in each year (in HT)

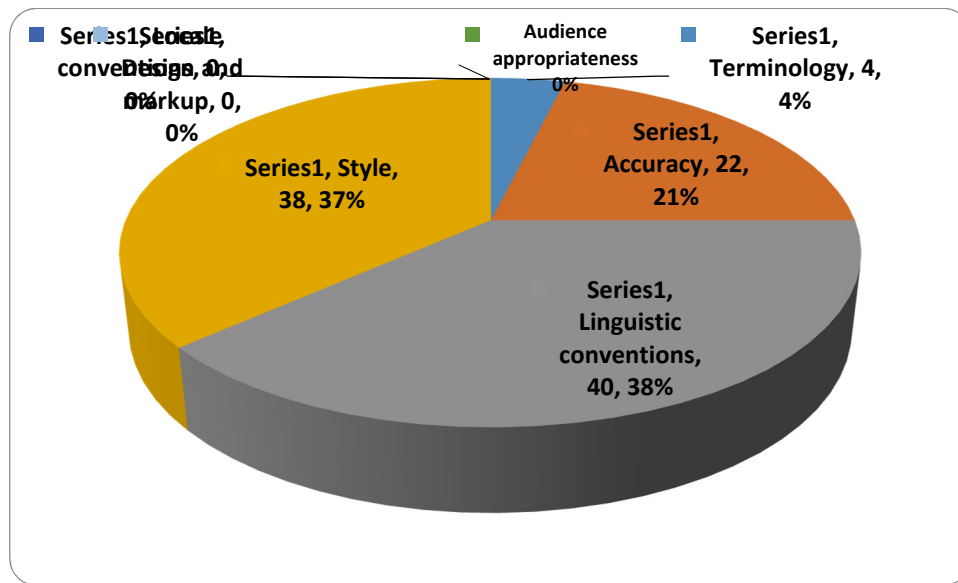
Error Dimensions \ Years	1 (1400)	2 (1401)	3 (1402)	4 (1403)	5 (1404)	Total
Terminology	1	1	2			4
Accuracy	7	4	7	2	2	22
Linguistic conventions	11	9	8	9	3	40
Style	12	12	8	2	4	38
Locale conventions						
Audience appropriateness						
Design and markup						
TOTAL	31	26	25	13	9	104
	30%	25%	24%	12%	9%	100%

The following line chart (4.3) is presented to make the general trend of error occurrence of in each year more tangible.

Figure 4.3 *Trend of top-level error dimensions occurred in 1400s (2021–2025)*



Based on the line chart data for the 1400s decade, the frequency of errors shows a consistent downward trend over time. Starting with 31 errors in 1400, the count decreases to 26 in 1401, followed by 25 in 1402. A more substantial drop occurs in 1403, with 13 errors, culminating in the lowest frequency of 9 errors in 1404. This represents a total reduction of 22 errors (71%) from the beginning to the end of the decade, indicating a marked improvement in translation quality over the five years. The total percentages of errors are displayed in Figure 4.4.

Figure 4.4 *Top-level error dimensions occurred in 1400s (2021–2025)*

Based on the data for the 1400s decade, the analysis of the seven top-level error dimensions across all five years (1400–1404) reveals that Linguistic conventions is the most frequent category, with 40 errors, accounting for 38% of all errors. Style follows closely with 38 errors (37%), while Accuracy recorded 22 errors (21%). Terminology contributed 4 errors (4%). The remaining dimensions—Locale conventions, Audience appropriateness, and Design and markup—registered no errors across the entire decade. These figures indicate that over the course of the 1400s, the primary and persistent challenges in translating sports science abstracts have been grammatical and mechanical language issues, closely followed by stylistic fluency, with accuracy concerns present but less dominant, while terminology, formatting, and cultural adaptation have been consistently well-managed.

4.1.3 Classification of errors

In table 4.4 frequency of the top level error dimensions are displayed in abstracts translated by human translator (from 2011 to 2025).

Table 4.4 *Frequency of errors occurred in each decade (in HT)*

Error Dimensions \ Decades	1390s	1400s	TOTAL
Terminology	17	4	21
Accuracy	40	22	62
Linguistic conventions	94	40	134
Style	12	38	50
TOTAL	163	104	267

Table 4.4 presents the total frequency of translation errors in English abstracts of theses across the specified period. In the 1390s (2011–2020), a total of 163 errors were recorded. This number decreased to 104 errors in the 1400s (2021–2025), reflecting a 36% decline compared to the previous decade. This consistent and accelerating decline in translation errors over the past 15 years suggests a notable improvement in the quality of English abstracts. Possible contributing factors may include enhanced translator training, greater awareness of academic writing standards, more rigorous editorial oversight, or the increasing use of NMT systems and AI tools. Moreover, a notable exception to the overall trend of error reduction is the Style category, which increased from 12 errors in the 1390s to 38 in the 1400s. This counterintuitive rise may partly reflect a discrepancy concerning the organizational style rather than an objective or clear cut linguistic error. At the same time, this trend could indicate a substantive challenge: translators may have concentrated on terminological and factual accuracy while devoting less attention to naturalness, fluency, adherence to TL academic conventions or organizational style. The simultaneous decline in other error categories and the rise in Style-related errors are consistent with a scenario in which overall competence improved, but stylistic polish did not keep pace.

4.2. Discussion

This section presents a qualitative analysis of selected instances from the corpus to explore the errors made by human translators and AI.

Source text (ST)	بطور مسلم داده هاي بيشتري براي تعريف نيمرخ پيكرسنجي ورزشكاران رشته هاي توپي، مقايسه ميان ورزش ها، ورزشكاران كشورهاي مختلف و سطوح مسابقات مورد نياز است.
Human translation (HT)	More data are certainly needed to define the anthropometric profile of martial arts athletes, the comparison of values among sports, athletes various countries and competition levels.
AI translation (AT)	Certainly, more data are needed to define the anthropometric profile of ball sport athletes, for comparisons between sports, athletes from different countries, and competition levels.

The HT contains a critical error. It translates "رشته های توپي" (ball sports) as "martial arts." This is a significant semantic error that completely changes the meaning of the sentence, misrepresenting the subject group of the study. The text "comparison of values among sports, athletes various countries and competition levels" is awkward and ungrammatical. It omits necessary prepositions, "athletes from various countries".

ST	ضمن اینکه تمرینات پلایومتریک شامل تماس های مکرر پا با زمین است، بنابراین احتمال بروز آسیب های اسکلتی-عضلانی به ویژه در زنان وجود دارد.
HT	Plyometric exercises contain frequent contact of foot with the land, thus athletes account a risk of musculoskeletal injuries specially in females.
AT	However, since plyometric exercises involve repeated foot contact with the ground, there is a potential risk of musculoskeletal injuries, particularly in women.

As can be seen in the HT, the Persian phrases "تماس ... پا با زمین" (foot contact with the ground/floor), and "احتمال بروز آسیب...وجود دارد" (athletes face/are at risk), are inaccurately translated as "contact of foot with the land" and "athletes account a risk of", respectively. The latter could also be translated as "there is a potential risk of" (AI's generated equivalent).

ST	جامعه آماری تحقیق حاضر را زنان سالمند ۶۰ تا ۷۴ سال شهر اصفهان تشکیل می دهد.
HT	The statistical population of the current research consists of elderly women aged with the age of 60 to 74 in Isfahan city.
AT	The statistical population consisted of elderly women aged 60 to 74 from Isfahan city.

The human translator awkwardly phrases the Persian text "زنان سالمند ۶۰ تا ۷۴ سال" (elderly women aged 60 to 74) as "elderly women aged with the age of 60 to 74," which is grammatically incorrect and redundant.

ST	هدف این پژوهش، تعیین اثر دوازده هفته تمرینات قدرتی منتخب در آب با خشکی بر سطوح ... می باشد.
HT	The aim of this research is the effect of twelve weeks of selected strength water and dryland training in water with dry land on the levels of
AT	The aim of this study is to determine the effect of twelve weeks of selected strength training in water versus on land on the levels of

In HT, the phrase "The aim of this research is the effect of..." could be revised as "The aim of this research was to investigate the effect of...". The HT also contains redundancy and is quite unnatural.

ST	درصد جمعیت افراد مسن دارای اضافه وزن و چاقی روند افزایشی دارد و در طول ۳۰ سال گذشته، میزان چاقی افراد مسن دو برابر شده است.
HT	The percentage of the elderly population with overweight and obesity is increasing and during in the last 30 years, the obesity rate of the elderly has doubled.

AT	The percentage of the elderly population with overweight and obesity is on the rise, with the rate of obesity among older adults having doubled over the past 30 years.
----	---

In HT, the phrase "during in the last 30 years" contains a redundant preposition. It could be revised as "during the last 30 years" or "in the last 30 years". The AT is more natural than the HT.

ST	... این تفاوت در مقایسه با گروه کنترل معنی دار بوده است.
HT	... this difference was meaningful in compare to control group.
AT	... this difference was significant for both experimental groups 1 and 2 compared to the control group.

In HT, the phrase "in compare to" could be revised as "compared to" or "in comparison with." The article "the" is also missing in the phrase "control group".

ST	سوختگی حادثه‌ای غم‌انگیز است که اختلالات زیادی در عملکرد جسمانی و سلامت روانی بیمار ایجاد می‌کند.
HT	Burn is one of the most horrendous traumas can suffer the body.
AT	Burns are a tragic incident that create numerous disruptions in patients' physical function and mental health.

The Persian text states "سوختگی حادثه‌ای غم‌انگیز است" (Burns are a tragic incident). Human translator has rendered it as "Burn is one of the most horrendous traumas can suffer the body," which adds intensity ("one of the most horrendous") not present in the Persian abstract. It also contains a grammatical error ("traumas can suffer the body") and it does not reflect natural academic English.

ST	لذا، هدف پژوهش حاضر، تاثیر تمرینات کراس فیت شامل: پرش، شنا، دراز نشست، ...
HT	Cross-fit exercises included: jump, swim, sit up, ...
AT	The Cross-Fit exercises included: jumps, push-ups, sit-ups, ...

The Persian lists exercises including "شنا" (push-ups). Human translator rendered this as "swim", which is completely incorrect. The error has been occurred due to the homonymous term "شنا", which can refer both to "push-up" and "swim". The hyphen in "sit-ups" is also missed in the HT.

ST	اضافه وزن با افزایش سن افزایش میابد و افراد بی‌تحرک‌تر میشوند.
----	--

HT	Overweight increases with age and sedentary people. They get wet.
AT	Weight tends to increase with age as individuals become more sedentary.

A very strange error that has occurred is the last sentence in HT: "They get wet ." After careful attention, we realized that the translator has made such an error by merely rendering "تر میشوند" in the last phrase in the Persian text "بی تحرک تر میشوند". A machine translation of early 2020 might make such an error. It does not, however, justify the ignorance of human translator to post edit his/her abstract.

ST	فاکتورهای مرتبط با آمادگی جسمانی (تعداد دراز نشست و حداکثر اکسیژن مصرفی)
HT	... factors related to physical fitness (long seating and maximum the consumed oxygen
AT	... physical fitness-related factors (number of sit-ups and VO2 max)

The Persian phrase "تعداد دراز نشست" (number of sit-ups) is rendered by human translator as "long seating". The translator has ignored the word "تعداد" (number) and the proposed equivalent for "دراز نشست" as "long seating" seems inaccurate—most probably due to the sheer reliance on the translation by a machine translation system (without being post-edited by a human expert). Additionally, it technically seems more appropriate to use "VO2 max" rather than "maximum the consumed oxygen" since brevity needs to be observed in abstracts. The mentioned phrase in the HT is also erroneous due to the misuse of the article "the".

ST	متغیرهای ترکیب بدن (... توده بدون چربی)، سطوح چربی های خون و... آثار تحمل وزن.
HT	Body composition variables (... and fat free mass), levels blood lipids ... weight bearing effects.
AT	Body composition variables (... and fat-free mass), blood lipid levels ... weight-bearing effects

The Persian phrase "توده بدون چربی" rendered by human translator as "fat free mass" should have been revised as "fat-free mass". The word order in the HT-generated phrase "levels blood lipids" is erroneous. It needs to be revised as "levels of blood lipid" or "blood lipid levels". Another hyphen is also missed in rendering the last phrase "آثار تحمل وزن" by human translator: "weight bearing effects".

ST	تمرینات مقاومتی منتخب در سنین نوجوانی اثربخشی یکسانی با تمرینات هوازی داشتند، و بنابراین این شیوه‌های تمرینی نیز در کودکان و نوجوانان مبتلا به دیابت نوع ۱ می‌تواند شیوه تمرینی مناسبی باشد.
HT	The efficacy of selected resistance training during adolescence is comparable to that of aerobic exercise, indicating that these training approaches may serve as suitable interventions for children and adolescents diagnosed with type 2 diabetes.
AT	Selected resistance training in adolescence had similar effectiveness to aerobic training, and therefore, these training methods can also be suitable exercise modalities for children and adolescents with type 1 diabetes.

A severe error has occurred in HT while rendering "دیابت نوع ۱" (type 1 diabetes) as "type 2 diabetes". Moreover, the AI-generated equivalent for the phrase "اثربخشی یکسانی با تمرینات هوازی داشتند" (had similar effectiveness to aerobic training) is semantically more accurate than the one produced by human translator: "is comparable to that of aerobic exercise".

Conclusions

The most significant translation challenges in the 1390s are concentrated in three primary categories: Linguistic Conventions, Accuracy, and Terminology. Other categories recorded zero errors across all years. The figures indicate that the primary challenges in translating the abstracts in this decade lie in maintaining grammatical accuracy, ensuring correct spelling, and avoiding omissions or mistranslations. While Linguistic Conventions and Accuracy require the most attention, lower-level issues such as Locale Conventions and formatting are perfectly managed, suggesting that the foundational aspects of localization and design are already well under control. During this decade, a noticeable decline in translation quality occurred, likely attributable to students' over-reliance on emerging MT technologies that had not yet achieved sophisticated capabilities.

Based on the analysis of the errors across the years 1400 (2021) to 1404 (2025), the data revealed that the most significant translation challenges are concentrated in two primary categories: Linguistic Conventions and Style. The results reveal that the primary challenges in translating the abstracts in this decade lie in achieving natural stylistic expression and maintaining grammatical accuracy. The years 1400 and 1401 emerge as the most problematic, with 31 and 26 total errors

respectively, driven largely by Language conventions issues. While Style and Linguistic Conventions require the most attention, lower-level issues such as Locale Conventions and formatting are rightly managed. It might suggest that the foundational aspects of localization and design are already well under control. In general, the number of errors decreased by the passage of time, which may indicate the utilization of new generations of MT-systems (based on NMT) or the use of AI by human translators. Moreover, based on the analysis of the errors detected in AI-generated translations, the data reveals optimal translation performance. It does not seem unjustifiable to assert that notable improvement in HT quality might be connected with the widespread presence of LLMs.

Overall, in the 1390s (2011–2020), a total of 163 errors were recorded. This number decreased to 104 errors in the 1400s (2021–2024), reflecting about a 36% decline compared to the previous decade. This accelerating decline in translation errors over the past 15 years suggests a notable improvement in the quality of English abstracts. Possible contributing factors may include: enhanced translator training, greater awareness of academic writing standards, or the increasing utilization of digital tools and AIs. This trend clearly reflects a positive shift toward higher clarity and accuracy in scholarly communication over time.

In line with Zandi's (2020) findings, which identified grammatical errors as the most frequent errors, the result of the present study revealed that, Linguistic conventions (including grammatical errors) consisted almost half of the total errors (38%). Moreover, based on the same researcher's findings, which specified addition and omission as the next highly frequent errors, the findings of the current study showed that about one third of the entire errors (about 21%) belonged to the category of Accuracy (including addition and omission errors).

The abstracts translated in the decade 1390s (2011 to 2020) consisted of translations most likely generated either by human on themselves or by human assisted MTs. The quality of abstract translations in this decade was tangibly lower than that generated in the 1400s. In the 1400s (2021 to 2025), with the emergence of the most advanced version of MT, namely NMT, followed by advanced AIs, we have also witnessed a reduction in the number of errors. Finally, while Li and Chen (2019) predicted that automated translation is 'unlikely' to replace human translation in the near future, it was found in the present study that AI-generated translations of abstracts in the field of physiology can be replaced by human-generated translations, while it may require merely minor revisions over a superficial post-editing process by an expert in the field. It, however, should be reminded that the researchers never

claim that human translators can totally be replaced by AI since our finding is based on limited number of abstracts from a single AI tool queried on a single occasion. In light of our small corpus, it was found that in a number of very specialized domains like abstracts of sports physiology, AI can produce very high quality translations which might be used after an even superficial post-editing process.

References

- Angelelli, C. (2009). Using a Rubric to Assess Translation Ability: Defining the Construct. In C. Angelelli and H. Jacobson (Eds.), *Testing and assessment in translation and interpreting studies. ATA Scholarly Monographs Series*. (pp. 13–47). John Benjamins.
- Asgari Vartooni, N. (2026). Book Review: Risk Management in Translation. *Íkala, Revista de Lenguaje y Cultura*, 31(2), 1–5. <https://doi.org/10.17533/udea.ikala.363898>
- Asgari Vartooni, N., Afrouz, M. & Minasian, V. (2026). Error Analysis of English Translations of Persian Sport Sciences Thesis Abstracts. *Translation and Interpreting Research*, 3(9), 1–18. doi: 10.22054/tir.2026.21086
- Bam, R. G. (2005). A Corpus-driven Quantitative Analysis of Translated and Original Texts in 'Die Burger' (Master's thesis). Stellenbosch University. <http://hdl.handle.net/10019.1/50242>
- Boynukara, E. (2017). On the Importance of Translation and Interpretation in Sports and the Reflections of Mistranslation. *Journal of Sport and Social Sciences*, 4(1), 1–6. <https://dergipark.org.tr/tr/pub/jsss/issue/33674/358372>
- Derakhshi, Z., & KhorramiFard, S. (2019). Quality Assessment of the Persian Translation of John Steinbeck's of Mice and Men Based on Waddington's Model of Translation: Application of Method A. *Journal of Language and Translation*, 9 (3), 29–41.
- Herdawan, D. (2020). An analysis on Indonesian-English abstract translation by Google Translate. *English Education: Jurnal Tadris Bahasa Inggris*, Vol. 13(2), 40–53.
- Jafarian, M. (2016). Qualitative Assessment of English Translation into Persian Abstract Iranian Students Dissertations based on the Role-oriented Pragmatic Model of the House (Master's thesis). Islamic Azad University Marvdasht branch. <https://ganj.irandoc.as.ir/>
- Li, H. & Chen, H. (2019). Human vs. AI: An Assessment of the Translation Quality Between Translators and Machine Translation. *International Journal of Translation, Interpretation, and Applied Linguistics (IJTIAL)*, 1(1), 1–12. <https://doi.org/10.4018/IJTIAL.2019010104>
- Löfgren, K. (2014). Terminology, Proper Nouns and Adaption: A Translation Study of Three Sports Texts (Bachelor's thesis), Linnéuniversitetet, Institutionenförspråk (SPR). <http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva>

- Lommel, A. R., Burchardt, A., & Uszkoreit, H. (2014). Multidimensional Quality Metrics (MQM): A Framework for Declaring and Describing Translation Quality Metrics. *Tradumàtica*, (12), 455–463. <https://doi.org/10.5565/rev/tradumatica.76>
- Mahdavi, E. (2017) A TQA of Persian Translations of M.A Theses Abstracts Based on Nord's Model: A Case Study of M.A Theses of Shahid Bahonar University Translation Students (Master's thesis) Kerman University. <https://ganj.irandoc.ac.ir/>
- Manipuspika, Y. S. (2021). The Effectiveness of Translation Quality Assessment (TQA) Models in the Translation of Indonesian Students. *International Journal of Linguistics, Literature and Translation*, 4(5), 287–297.
- Munawir, M. (2020). The Kinds of Translation Error Made by the Students in Writing Abstract of the Theses and Dissertations. *Sang Pencerah: Jurnal Ilmiah Universitas Muhammadiyah Buton*, 6(2), 58–66. <https://doi.org/10.35326/pencerah.v6i2.723>
- Napitupulu, S. (2017). Analyzing Indonesian-English Abstracts Translation in View of Translation Errors by Google Translate. *International Journal of English Language and Linguistics Research*, 5 (2), 15–23.
- Terzi, C., & Arslantürk, Y., (2014). An Analysis of Dissertation Abstracts In Terms of Translation Errors and Academic Discourse. *International Journal of English Language and Translation Studies*, 2(4), 1–11.
- Yuan, B. (2019). On Sports English Translation: A Perspective from Linguistic Theories of Translation. *Journal of Chengdu Sport University*, 45(1): 73–79. DOI: [10.15942/j.jcsu.2019.01.011](https://doi.org/10.15942/j.jcsu.2019.01.011)
- Zandi, M. A. (2020). A Translation Quality Assessment of Thesis and Article Abstracts Translated from Persian into English in the Field of Psychology (Master's Thesis). University of Isfahan.