*Salud, Ciencia y Tecnología*. 2025; 5:1426
doi: 10.56294/saludcyt20251426

ORIGINAL

# Bridging the Language Gap: Enhancing Academic Performance of Non-Native Students with AI-Powered Translation

## Cerrar la brecha lingüística: mejorar el rendimiento académico de estudiantes no nativos con traducción impulsada por IA

Shee Mun Yong[1], Ah Kiaw Phan[1]

[1]University of Wollongong Malaysia, School of Business. Shah Alam, Malaysia.

**Cite as**: Yong SM, Phan AK, Bridging the Language Gap: Enhancing Academic Performance of Non-Native Students with AI-Powered Translation. 2025; 5:1426. https://doi.org/10.56294/saludcyt20251426

**Submitted:** 23-06-2024 **Revised:** 02-10-2024 **Accepted:** 28-05-2025 **Published:** 29-05-2025

**Editor:** Prof. Dr. William Castillo-González

**Corresponding Author:** Shee Mun Yong

## ABSTRACT

**Introduction:** disparity between language ability of non-native students and the medium of instruction used in host country education institutions can limit non-native students' (NNES) ability to be fully engaged in learning and if not addressed can pose serious ramifications on affected students' psychological wellbeing leading to poor academic performance and unwanted dropout.
**Objective:** this study uses empirical evidence to establish the effectiveness of AI neural machine translators in enhancing non-native students' academic performance through improved learner-content and learner-instructor interactions.
**Method:** convergent parallel mixed method was employed over two trials simultaneously collected quantitative and qualitative data from samples that were segregated into similar proportions of control and experiment sub-population strata according to participants' language ability. Qualitative semi-structured interviews were conducted during the diagnostic and summative stage complemented by weekly formative quantitative assessment over a span of 10 weeks per trial.
**Results:** findings from Cohen's d's effect size change inferred NNES with low language ability benefited the most. However, lesser effects were found on NNES with high language ability.
**Conclusions:** while NMT shows promise in enhancing learning interactions educators should exercise discretion to avoid deleterious effect on host country students and NNES with higher language abilities.

**Keywords:** Neural Machine Translator; Convergent Parallel Mixed Method; Cohen's d; Effect Size Change.

## RESUMEN

**Introducción:** la disparidad entre la capacidad lingüística de los estudiantes no nativos y el medio de instrucción utilizado en las instituciones educativas del país anfitrión puede limitar la capacidad de los estudiantes no nativos (NNES) para participar plenamente en el aprendizaje y, si no se aborda, puede plantear graves ramificaciones para los estudiantes afectados. 'Bienestar psicológico que conduce a un bajo rendimiento académico y a un abandono escolar no deseado.
**Objetivo:** este estudio utiliza evidencia empírica para establecer la efectividad de los traductores automáticos neuronales de IA para mejorar el rendimiento académico de los estudiantes no nativos a través de mejores interacciones entre el alumno y el instructor.
**Método:** se empleó un método mixto paralelo convergente en dos ensayos que recopilaron simultáneamente datos cuantitativos y cualitativos de muestras que se segregaron en proporciones similares de estratos de subpoblación de control y experimento de acuerdo con la capacidad lingüística de los participantes. Se realizaron entrevistas cualitativas semiestructuradas durante la etapa de diagnóstico y sumativa, complementadas con una evaluación cuantitativa formativa semanal durante un lapso de 10 semanas por prueba.

**Resultados:** los resultados del cambio en el tamaño del efecto de la d de Cohen infirieron que las NNES con baja capacidad lingüística fueron las que más se beneficiaron. Sin embargo, se encontraron efectos menores en NNES con alta capacidad lingüística.
**Conclusiones:** si bien la NMT es prometedora para mejorar las interacciones de aprendizaje, los educadores deben actuar con discreción para evitar efectos nocivos en los estudiantes del país anfitrión y en las NNES con mayores habilidades lingüísticas.

**Palabras clave:** Traductor Automático Neuronal; Método Mixto Paralelo Convergente; D De Cohen; Cambio de Tamaño del Efecto.

## INTRODUCTION

With Malaysia transitioned to post Covid-19 endemic phase and the reopening up of its borders to international travelers, the education sector has mostly returned to classroom study from online learning model implemented during the pandemic. Students, especially those who enrolled in education institutions at the onset of the pandemic, have no experience of classroom or campus learning, suffered anxiety and psychological stress due to different learning pedagogies used during and after the pandemic.[1] This phenomenon is exacerbated for international students from non-English speaking countries such as China. Malaysia has seen a sharp rise of more than 150 % of Chinese students due to geopolitical shift.[2] Particularly, there are approximately 39 000 Chinese students, majority of whom enrolled in private higher education institutions.[3] As English is considered lingua franca predominantly among urban dwellers and business communities in Malaysia, the medium of instruction in most private higher learning institutions (HLI) is English, while public HLIs use a combination of English and Malay language. This language disparity causes difficulties for Non-Native Non-English-Speaking Students (NNES), who struggle to understand the subject taught as the medium of instruction, learning materials and assessments are all in language stipulated by their learning institution, leading to low confidence and engagement in learning.[4] Consequently, they are unwilling to communicate in class due to fear of negative evaluations and being judged unfavorably by others. If left unattended, it could negatively affect student's social function, academic achievement, or even physical health.[5] Past studies have attested to the correlation between language proficiency and academic performance.[6] Students without adequate language proficiency levels are more likely to face difficulty in understanding subject matters, while those proficient in instruction language perform better academically.[7] It is therefore essential for schools and educators to prioritize the social, emotional, and academic learning of NNES to improve their psychological wellbeing and ease their adaptation to on-campus life. This challenge could be alleviated by translating the medium of instruction and all learning materials into the NNES native language; however, outcome was not consistent. Past studies posited that grammar-translation methods only promote passive learning[8] whilst other studies have highlighted the benefits of speech-enabled language translation in improving students' attention and meditation.[9] This raises the question about the effectiveness of using neural machine translator as pedagogical tool to enhance NNES's psychological wellbeing and academic performance that form the aim of this research.

### Literature Review and Hypotheses Development

The advent of the 4th Industrial Revolution introduced Neural Machine Translation (NMT), which uses neural network models to build statistical models using human cognition that serves as communicational computer-human interaction in translating written or spoken words from one language to another.[10] NMT recognizes the source patterns and uses them to determine a context-based interpretation as likelihood of the word sequence. Recent developments have shown that utilizing these deep learning artificial neural networks such as deep neural networks (DNN), convolutional neural networks (CNN), and recurrent neural networks (RNN) have significantly improved language translation when applied to Natural Language Process (NLP).[11]

### Machine Translation in Education and Learner's Cognitive Abilities

AI's impact on education has been extensively researched, uncovering opportunities for developing cognitive intelligence and effective student learning that educators have taken advantage of to improve teaching methods using adaptive learning platforms (ALP). NMT technology as part of AI-enabled applications has made major advances in language translation, providing higher quality and more human-sounding output.[12] Companies such as Microsoft and Google leverage on these deep neural networks to enhance text and speech application program interface (APIs) powering high-quality speech-to-text (STT), text-to-text (TTT), and text-to-speech (TTS) translation.[13] Although NMT can be an efficient tool suitable for productive (speaking and writing) and receptive language skills (reading and listening) of non-native speakers,[14] past studies have yielded mixed results. Despite some findings showing improvement in participants' reading comprehension,[15]

writing skill,[16] traditional face-to-face, integrative, and web-based learning,[17] others showed inconsistency in communicative language teaching and handling of socio-culturally complex text.[18] These mixed results can partially be explained from the use of different types and versions of machine translation technology that necessitates further exploration into using deep neural networks enabled machine translation for language learning research.

Machine translation has been proven in developing learner's cognitive abilities such as communication competence[19] reading comprehension,[15] writing qualities[20] and positive attitudes.[21] User-friendly low-cost online machine translation tools such as Google Translate and Microsoft Translator are well known for utilizing NMT technology to facilitate learners' interaction,[22,23] the advancement of NMT output quality, such as post-editing of automatic translated texts can improve command in both mother and second language.[24] Despite Google Translate's relatively high rate of grammatical errors, its accuracy meets the minimum level of accuracy required for university entrance.[25] This suggests that it is not just using translation tools but more importantly how one utilizes these tools for fit-for-purpose applications.[26] For instance, dissemination of published information would require human assistance, while assimilation and interchange of information for non-specialist general audience does not require high-level human intervention.[27] In other words, students should focus on understanding the overall context of information, instead of word-for-word translation, in fostering better meaning, especially in text that has complex lexical structure.

**Effectiveness of Machine Translation in Achieving Learning Outcomes**

With the advancement in NMT, improved accuracy has considerably made it more attractive for educational settings where high-level translation accuracy is not required. Institutions like Massachusetts Institute of Technology, and Johns Hopkins Bloomberg School of Public Health use translation in their online courses for foreign students.[28] Despite inconsistent quality, students studying non-linguistic subjects derived greater positive impacts from translation of difficult words, sentences, and paragraphs.[29] These studies provide testimonials of the effectiveness in using text-to-text (TTT) translation to foster learner-content interaction and comprehension of complex concepts. However, not much is known about the effectiveness of a real time speech-to-text (STT) NMT in enhancing learner-instructor interaction. The lack of post-editing process, speaker's accent, grammar, and nuances in sentences have contributed to reduced NMT accuracy. It raises the question of how these inaccuracies affect learner-content and learner-instructor interaction and hence, their learning outcome and whether using NMT would degrade academic performance of NNES. In summary, this research study aims to answer the following questions:

RQ1: what effect does TTT NMT have in enhancing NNES learner-content interaction?

RQ2: would using STT NMT in class to facilitate lesson delivery enhance NNES learner-instructor interaction?

RQ3: would the combination of enhanced learner-content and learner-instructor interactions lead to an improved academic performance?

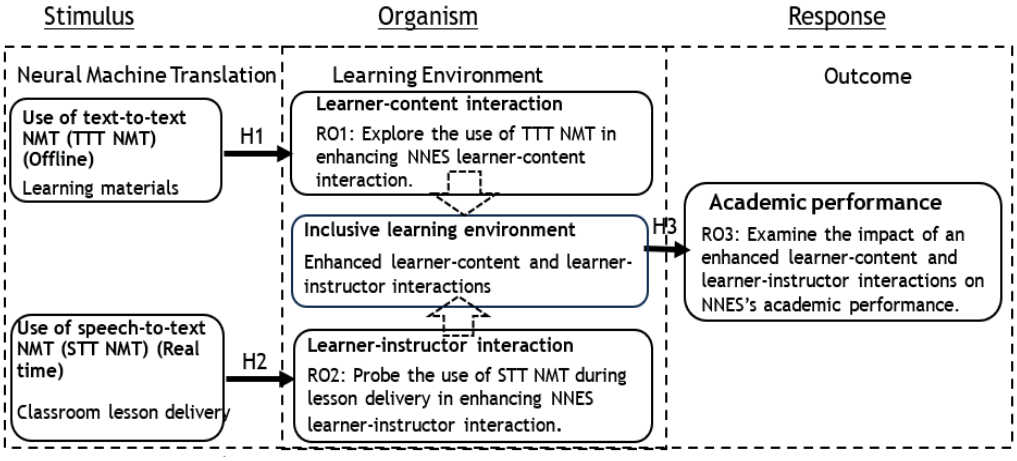


**Figure 1.** Conceptual framework and underpinning theory

From these core questions, research objectives were formed and depicted in the conceptual framework (figure 1), supported by Stimulus-Organism-Response (SOR) theory. This theory evaluates the effectiveness of NMT in formal learning context by interpreting the application of TTT and STT NMT as stimulus (S) to motivate cognitive transfer knowledge. This creates an inclusive and conducive learning environment that leads to enhanced organismic (O) learner-content and learner-instructor interaction. The outcomes are positive responses (R) in the form of improved academic performance. Consequently, the formulated hypotheses are:

H1: TTT NMT promotes learner-content interaction that results in significant linkage between the two variables.

H2: STT NMT as complementary pedagogical tool would enhance NNES learner-instructor interaction.

H3: academic performance of NNES is positively impacted by enhanced learner-content and learner-instructor interaction.

## METHOD

This study utilizes a convergent parallel mixed method design to simultaneously collect quantitative and qualitative data. The data strands are then analyzed separately with outcomes integrated and triangulated by interpreting and validating the quantitative outcome with findings from qualitative analysis[30] (figure 2).

### Population and Sample Size

G*power software was used to determine the study sample size because of its ability to generate optimum sample size for statistical T-Test between control and experiment group, minimizing the possibility of Type I and Type II errors. To achieve 80 % power for a two-tail test with equal group size, standard alpha (α) value of 0,08, beta (ß) value of 0,5 and power value of 0,8 were applied that generated a sample size of 37 participants per group.

### Research Approach

The study was conducted over two 14-week semesters of classroom study. Each semester covers four main stages: Diagnostics, Formative, and Summative assessment, following the approach by Chufama et al.[31] and a final Integration stage added to interpret findings from these assessments (figure 2).

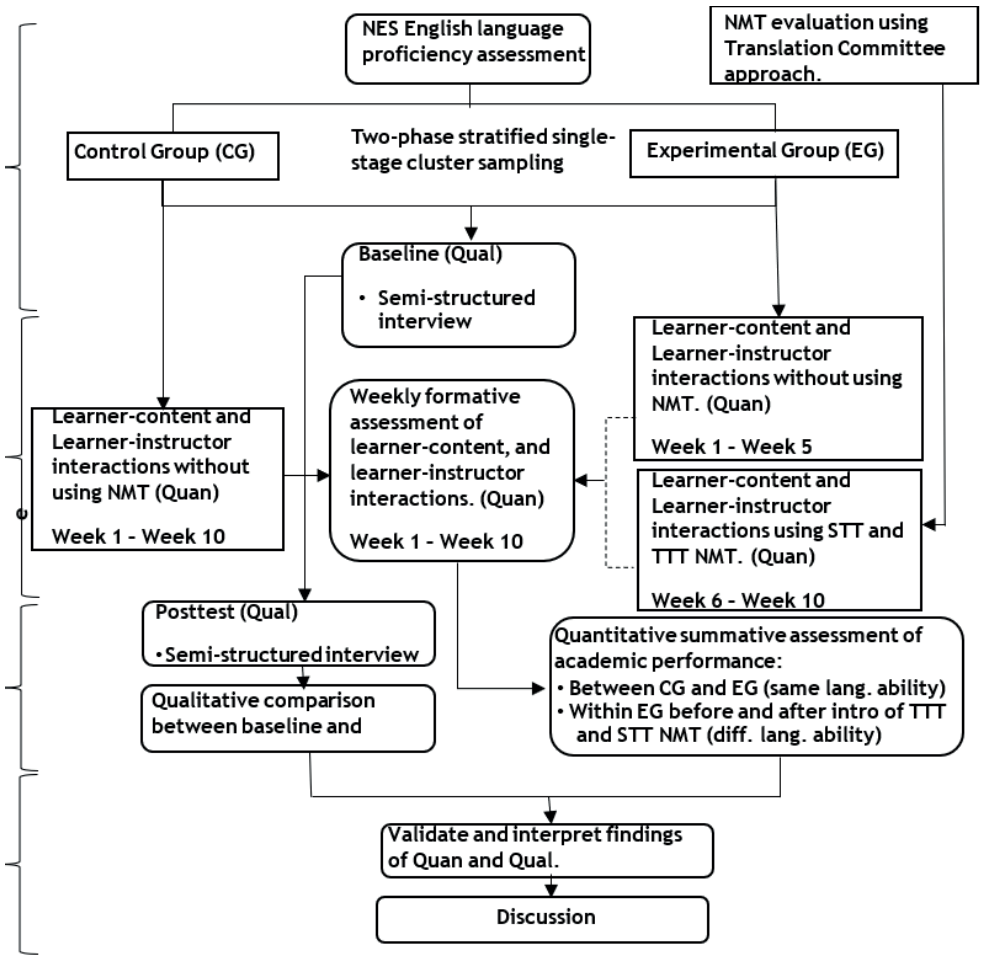


**Figure 2.** Convergent-parallel mixed method

*Diagnostics Assessment*

A two-phase stratified single-stage cluster sampling approach[32] was carried out in the first week. For the first phase, participants were divided into control and experimental groups based on scores of English proficiency tests. Data was drawn from each subject within these mutually exclusive clusters in the second phase to ensure representativeness and avoid sample bias. Baseline data was formed through qualitative semi-structured interview for both groups to measure the effect of enhanced learner-content and learner-instructor

https://doi.org/10.56294/saludcyt20251426

interactions on NNES. A three-step process of forward translation, consensus, and reconciliation of translation committee[33] was adopted leading to Microsoft Translator chosen as NMT tool for TTT and STT, due to its ease of use in MS PPT real time caption and its flexible features.

*Formative Assessment*

For the first five weeks, both control and experimental groups were assessed weekly using quizzes as academic metrics in measuring their academic performance. TTT and STT NMT were introduced in the following five weeks to the experiment group whilst the control group continues to receive the same pedagogical approach. Formative assessment for both groups continued until the end of semester.

*Summative Assessment*

Post-test semi-structured interviews were conducted, and the data was compared with the baseline data. Additionally, the weekly formative assessments were analysed to measure cognitive transfer between experimental and control groups for NNES of the same language ability and within experimental and control groups for NNES of different language ability before and after the introduction of TTT and STT NMT.

*Integration Stage*

The findings from comparative analyses of the two observed groups were collated and interpreted to explain the outcomes that connect the strands of data from qualitative interview, and empirical results of NNES'academic performance.

## RESULTS

The first and second trials involved 72 NNES undergrads and 72 NNES postgrads respectively. Multiple trials were conducted to ensure the reliability of results and avoid the influence of other non-related events. Additionally, testing at different educational levels allowed researchers to monitor the consistency of the result across various levels of education. Participants were divided into three groups based on their TOEFL and IELTS scores. Chinese NNES were chosen for this study because they are the largest single nationality from a non-English speaking country enrolled in the institution of which this study was conducted. The first group, Low Language Ability (LLA), comprised of NNES who can understand and communicate in basic expressions. The second group, Medium Language Ability (MLA), is made up of independent NNES who could interact with some fluency and spontaneity. The final group, High Language Ability (HLA), comprised of NNES who could summarize information and arguments coherently and concisely.[34] Each group was further segregated into control and experiment sub-population strata of equal size, upon which samples were drawn (table 1).

**Table 1.** Stratified single-stage cluster sampling

| **NNES Population** | **72 (Undergrad NNES)** | | **72 (Postgrad NNES)** | |
|---|---|---|---|---|
| Stratified single-stage cluster sampling | Control group | Experiment group | Control group | Experiment group |
| LLA (A1/A2 basic user) | 12 | 12 | 12 | 12 |
| MLA (B1/B2 Independent user) | 12 | 12 | 12 | 12 |
| HLA (C1/C2 Proficient user) | 12 | 12 | 12 | 12 |

**Qualitative Analysis**

In Phase 1, 72 undergrad NNES (Experimental group (N = 36, Control group (N= 36)) and in Phase 2, 72 postgrad NNES (Experimental group (N = 36), Control group (N= 36)) participated in semi-structured interviews. These interviews were conducted in the first and tenth week of each semester, using the same interview guide which focused on the three areas of 1. accuracy, clarity, and vocabulary, 2. facilitation of class activities/tutorial questions, and 3. enhancement of learner-content interaction, perceived impact on class performance.[35]

*Effectiveness of TTT NMT in Enhancing Learner-Content Interaction*

Hypothesis (H1) was tested by comparing the experiences of three language ability levels of NNES experimental and control groups before and after TTT NMT implementation. Even though both groups held initial favourable perceptions of the tool they nevertheless expressed cautious optimism, highlighting potential issues with accuracy, clarity, and vocabulary, especially for nuanced language and technical terms. Notably, students with HLA hold a general uncertainty concerning the TTT NMT's potential benefits and effectiveness. By the end of the trial, perception shifted from pre-experiment optimism to a more critical stance, especially among HLA. Although HLA continues to exhibit cautious optimism, they found it less useful due to occasional inaccuracies in

translation and insufficient handling of specialized terminologies. This concern was valid as they were already comfortable with the learning materials prepared in the medium of instruction. On the other hand, LLA and MLA NNES maintained a positive view on the effectiveness of TTT NMT, as they experienced notable improvements in understanding the class materials. To them, TTT NMT was seen as a tool that effectively reduced learning barriers and enhanced the efficiency of completing class activities. Hence, the findings support hypothesis (H1) that TTT NMT enhances NNES learner-content interaction and improves NNES' understanding of learning materials.

*Effectiveness of STT NMT in Enhancing Learner-Instructor Interaction*.

Similarly, hypothesis (H2) was tested by comparing experiences before and after implementation of STT NMT. From an initial favourable expectation, a gap began to develop between expectation and actual experience particularly among HLA NNES who anticipated a higher level of precision from the translation technology. The inconsistency in translation accuracies hindered the effectiveness of communication causing HLA NNES to revert to conventional classroom learner-instructor interaction and used STT NMT only when they faced difficulty in understanding the instructor's expressions. In contrast, both LLA and MLA NNES undergrads reported improvement in understanding lessons taught, with LLA NNES benefiting the most. However, mixed results were reported among postgrad MLA NNES, contrasting better lesson understanding with uncertainties on inaccurate technical terminology translation hence deducing a partial supportability of hypothesis (H2) in using STT NMT as complementary pedagogical tool.

**Quantitative Analysis**

*Effect of Inclusive Learning Environment on Academic Performance*.

The hypothesis (H3) is validated using a quantitative study carried out in the form of quizzes that were administered as metrics to measure NNES's academic performance. To reduce the influence of extraneous and confounding variables, NNES were asked to complete the assessment at the end of each lesson. During the first trial period, T-Test performed between experiment and control NNES groups (table 2) showed insignificant finding (sig. > 0,05) for all the tested groups.

**Table 2.** T-test and Effect Size Change (ESC)

| | | (Wk2 - Wk5) | | | | | | (Wk7 - Wk10) | | | | | | Effect |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Exp. Grp Mean | Std. Dev. | Ctl. Grp Mean | Std. Dev. | T-Test | d | Exp. Grp Mean | Std. Dev. | Ctl. Grp Mean | Std. Dev. | T-Test | d | Size Chg. (ESC) |
| Undergrads (UG) | LLA | 2,100 | 1,137 | 2,400 | 1,671 | 0,673 | 0,210 | 3,667 | 0,720 | 2,517 | 0,835 | 0,030 | 1,475 | 1,265 |
| Undergrads (UG) | MLA | 2,600 | 0,966 | 2,950 | 1,048 | 0,560 | 0,347 | 4,183 | 0,328 | 3,283 | 0,458 | 0,038 | 1,259 | 0,912 |
| Undergrads (UG) | HLA | 3,517 | 0,519 | 4,050 | 0,896 | 0,240 | 0,730 | 4,033 | 0,935 | 4,417 | 0,685 | 0,440 | 0,468 | -0,262 |
| PostGrad (PG) | LLA | 3,017 | 0,564 | 3,158 | 0,304 | 0,650 | 0,311 | 4,217 | 0,449 | 3,233 | 0,301 | 0,001 | 2,574 | 2,263 |
| PostGrad (PG) | MLA | 3,800 | 0,415 | 3,783 | 0,611 | 0,957 | 0,033 | 4,433 | 0,437 | 3,933 | 0,320 | 0,040 | 1,306 | 1,273 |
| PostGrad (PG) | HLA | 4,350 | 1,098 | 4,183 | 0,431 | 0,737 | 0,200 | 4,767 | 0,723 | 4,467 | 0,677 | 0,475 | 0,428 | 0,228 |

**Note:** Cohen's d effect size: small effect (d = 0,2), medium effect (d = 0,5), large effect (d = 0,8)

However, when the periods before and after the introduction of NMT were compared, LLA and MLA NNES for both under and postgrads showed notable improvement changing from an insignificant mean value difference (sig. > 0,05) to a large mean value difference (sig. < 0,05). In contrast, improvement for HLA NNES is marginal and insignificant (sig. > 0,05). As sig. in T-Test indicates an effect exists but not the magnitude of a change, Cohen's d (d) was used to give a better interpretation of the results with effect size classified as small (d = 0,2), medium (d=0,5), and large (d > 0,8).[36] The results clearly indicate that LLA NNES has the highest effect size at the end of trial period followed by MLA NNES with marginal change for HLA NNES. Unfortunately, effect size shows only the outcome of comparison between experiment and control groups for each of the periods but lack details on the effect of change as students transitioned from the first period to the second period. To study the change effect, "Effect Size Change (ESC)" was computed based on the difference in effect size (d) between the periods before and after introducing NMT. From the analysis of ESC, the statistics show positive

academic performance improvement for both LLA and MLA NNES groups with LLA NNES in the experiment group having the highest effect size change and thus considered the most improved group. Unsurprisingly, HLA NNES was the least affected. There was even a slight degradation in performance for UG HLA NNES. The finding infers the effectiveness of NMT in fostering a more inclusive learning environment leading to better comprehension of lessons taught among NNES with lower language proficiency. Although HLA NNES has benefited to a certain extent, it is not large enough to make a difference in their overall cognitive ability.

However, comparison between experiment and control groups made between experiment and control groups of NNES with same language proficiency does not tell the progress in relation to other groups which begs the question of whether NMT can help lower language ability NNES close the gap with NNES of higher language ability. To achieve this objective, SPSS One-Way Anova supported by Tukey's post hoc test was undertaken to statistically compare the mean differences among the three NNES groups (table 3). Based on similar arguments that effect size alone cannot measure NNES's transition cognitive transfer process necessitates the computation of Effect Size Change (ESC) (table 4).

**Table 3.** Anova and Effect Size

| | | Experimental Groups | | | | | | | | Control Groups | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | (Wk2 - Wk5) | | | | (Wk7 - Wk10) | | | | (Wk2 - Wk5) | | | | (Wk7 - Wk10) | | | |
| | | Tukey HSD sig. | | d | | Tukey HSD sig. | | d | | Tukey HSD sig. | | d | | Tukey HSD sig. | | d | |
| | | LLA | MLA | LLA | MLA | LLA | MLA | LLA | MLA | LLA | MLA | LLA | MLA | LLA | MLA | LLA | MLA |
| UG | MLA | 0,618 | | 0,474 | | 0,537 | | 0,922 | | 0,731 | | 0,394 | | 0,157 | | 1,137 | |
| UG | HLA | 0,042 | 0,223 | 1,603 | 1,183 | 0,726 | 0,947 | 0,439 | 0,214 | 0,059 | 0,309 | 1,231 | 1,128 | 0,001 | 0,028 | 2,488 | 1,946 |
| PG | MLA | 0,202 | | 0,564 | | 0,779 | | 0,488 | | 0,083 | | 1,295 | | 0,078 | | 2,253 | |
| PG | HLA | 0,020 | 0,434 | 1,527 | 0,415 | 0,229 | 0,561 | 0,914 | 0,559 | 0,005 | 0,325 | 2,748 | 0,757 | 0,001 | 0,120 | 2,355 | 1,009 |

**Table 4.** Effect Size Change (ESC)

| | | Experimental Groups | | Control Groups | |
|---|---|---|---|---|---|
| | | LLA | MLA | LLA | MLA |
| UG | MLA | 0,448 | | 0,743 | |
| UG | HLA | -1,164 | -0,969 | 1,257 | 0,818 |
| PG | MLA | -0,076 | | 0,958 | |
| PG | HLA | -0,613 | 0,144 | -0,393 | 0,252 |

The significant (sig. < 0,05) effect size between LLA and HLA NNES (LLA→HLA) groups at the beginning of trial period for both undergrad and postgrads indicates a large disparity in academic performance. By the end of the trial, Tukey results show an insignificant result (sig. > 0,05) which implies that LLA NNES has caught up with the academic performance of HLA NNES. This outcome is supported by the control group that continue to show a large disparity (sig. > 0,05) in academic performance between LLA and HLA NNES groups. In the case of LLA→MLA, there were mixed results. Although Experiment PG LLA → PG MLA showed a slight improvement (ESC: -0,076), UG LLA → UG MLA (ESC: 0,448) showed a worse outcome. However, it is worth noting that Control UG LLA → UG MLA (ESC: 0,743) showed the worst outcome inferring the positive effect of NMT. In contrast Experiment PG MLA → PG HLA showed a slight decline (ESC: 0,144) as opposed to an improvement for UG MLA → UG HLA (ESC: -0,969). This finding is consistent with the qualitative outcome of mixed results with UG MLA NNES perception of an improvement in understanding of lesson taught and a less favourable PG MLA NNES perception.

Overall, there is a general improvement in academic performance compared to the control group. For instance, the ESC value (Experiment UG LLA→ UG HLA: -1,164) showed an improvement in LLA NNES's academic performance in relation to HLA NNES. In contrast, ESC value of the Control group (UG LLA → UG HLA: 1,257) shows degradation in their academic performance. The outcome supports Hypothesis (H3) which attests to the effectiveness of NMT in creating a conducive and inclusive learning environment leading to better academic performance.

## DISCUSSION

As a recap, the objectives for this study are:

RO1: explore the use of TTT NMT in enhancing NNES learner-content interaction.

RO2: probe the use of STT NMT during lesson delivery in enhancing NNES learner-instructor interaction.

RO3: examine the impact of enhanced learner-content and learner-instructor interactions on NNES's academic performance.

RO1 was achieved with findings supporting hypothesis (H1), such that TTT NMT enhanced NNES learner-content interaction among NNES, particularly for those with low language abilities. They demonstrated substantial improvement in understanding class materials and engaging in academic activities, while those with higher language abilities found occasional inaccuracies and limitations of TTT NMT in handling special terminologies but did not find these issues disruptive in their learning. These findings align with previous research,[15,22,29] highlighting the effectiveness of NMT in enhancing learner's reading comprehension and facilitating the learner-content interactions. These findings are plausibly attributed to two reasons. First, the gap in language proficiency and the medium of instruction is greater in LLA NNES, making it harder for them to comprehend the learning materials. TTT NMT bridges this gap by making the materials more accessible and enhancing the learner-content interaction. Second, comprehending learning materials in non-native language increases learners' cognitive load particularly for those with lower language abilities. TTT NMT reduces this load, allowing LLA NNES to understand the context of the message, with the correct nuances and structures being less critical. HLA NNES, who possess better language fluency, are more critical of language structures and less tolerant of translation inaccuracies. Nevertheless, TTT NMT has proven effective in alleviating language barriers and enhancing learning experience, allowing NNES to adapt better in a new learning environment.

In undertaking the second objective (RO2), the outcome of this study partially supports hypothesis (H2) on the effectiveness of STT NMT in enhancing NNES learner-instructor interaction. Despite findings showing STT NMT as a versatile pedagogical tool in generating understanding and easing communication among LLA NNES, HLA NNES found the tool a distraction. The findings corroborate somewhat with previous research.[37,38] Whilst these past studies determined the positive impact of machine translation, particularly STT NMT in educational settings, this study establishes a diminishing value of the translation tool in relation to an increase in language ability. It signifies an inverse correlation between NNES's language ability and the benefit from using NMT as pedagogical tool. This assertion is attributed to LLA NNES struggle with the pace of instruction during lessons. STT NMT real-time translation helped overcome this difficulty, boosting their confidence and willingness to participate in class discussion, leading to an enhanced learner-instructor interaction. In contrast, HLA NNES found that the instructors' use of complex language structures, pace of delivery, and accent exacerbated STT NMT translation inaccuracy and distorted the cognitive transfer process.

For the third objective (RO3), analysis of formative assessment reveals that LLA NNES benefited the most from the use of NMT. This is attributed to two reasons. First, STT NMT has improved significantly in translation quality and fluency, making the translations adequate for these struggling NNES to better comprehend class lessons. Second, their confidence and motivation were further boosted by TTT NMT enhanced learner-content interaction. Contrary to studies[39] that more advanced L2 foreign language students benefited the most from machine translation, this study found that LLA NNES demonstrated the most improvement in academic performance. While the former assessed students' writing skills, this study attests the amelioration to greater understanding and comprehension of the topic taught. In this regard, the poor performance among NNES may not be due to their aptitude but rather the cognitive ability of an individual in comprehending the medium of instruction. This supports the third hypothesis (H3) that enhanced leaner-instructor and learner-content interaction facilitated by NMT enables LLA students to perform at parity with their higher language ability peers, challenging the notion that students with low language ability are academically poor students. It is therefore unsurprising that higher language proficient students do not benefit much from NMT, as their grasp of the medium of instruction is adequate to understand the class lessons and not much is needed for this supplementary aid. For these students, knowledge transfer and cognitive ability are no different from the host country students who rely solely on the medium of instruction in the learning process.

This study has found that NMT is not suitable for all levels of NNES. Educators should conduct a diagnostic assessment of students'language ability level prior to using it to prevent possible deleterious effects on host country students and NNES with higher language proficiency.

**Theoretical and Practical Implications**

While the results demonstrate the positive impact of TTT NMT on learner-content interaction and to a certain extent STT NMT, it is crucial to recognize the ongoing challenges in NMT quality and accuracy despite advancement made. First, multimodal and cross-modal machine translation requires large data corpora, which would be ineffective for language with low resources. Second, evolution of languages exacerbated by rapid sociocultural and technological developments make it harder for researchers to keep up with ever-changing linguistic knowledge and patterns. Third, lexical ambiguity, accent, pace, and complex grammatical structure in STT NMT are added to these challenges.

Even though TTT NMT shows greater reliability, consistency, and accuracy in translation, it cannot be used as an STT approach, which requires automated speech recognition (ASR) transcription before translation. Within STT, there is a trade-off between latency and quality of the translation. To improve the quality, STT tools often transcribe what it perceives as full sentence using punctuation as delimiter of that assumption; thereby increases STT latency and often result in discorded synchronization between instructor gesture, slides and translated text. Dynamic transcription, where ASR transcribes incrementally as new words are received without waiting for the end of a sentence,[40] can alleviate some latency issues but does not address issues of lexical ambiguity, accent and complex grammatical structure. In this regard, educators must be aware of these limitations and contribute to improving translation quality by speaking clearly, avoiding technical jargon, using simple grammar structure, reducing sound interference from the surrounding, minimizing classroom noise from within the classroom and using longer pauses as delimiter of end sentence.

The presence of STT translations can be distracting for other classmates, as the projected view on a single monitor is constricted by the presence of STT targeted language, making it smaller and more difficult to see, especially the smaller prints. This can be a distraction for host country students and HLA NNES, thereby reducing knowledge transfer. To mitigate this issue, the original slides and TTT translated slides, together with STT, can be displayed on separate projector screens. Multi-monitor software such as OfficeOne Powershow, allows audience to view their preferred screen without being distracted by smaller prints or translated text.

From the discussion, it is clear that Stimulus-Organism-Response (SOR) used to underpin the effectiveness of NMT in formal learning context only addresses the process aspects of applying TTT and STT NMT as stimulus (S) in enhancing organismic (O) learner-content and learner-instructor interaction experience and improving academic performance. It does not address the psychological part of NNES wellbeing, which is crucial for academic performance. Past studies [41] have highlighted the correlation between wellbeing and students'academic performance. For LLA NNES, NMT serves as hygiene factor in Herzberg two-factor theory and the improved cognitive ability provides the stimulus and intrinsic motivation to achieve satisfactory academic performance.[42] Study [9] have recognized that STT NMT enhances perceived academic emotions and learning satisfaction among LLA students. Having an inclusive learning environment facilitated by NMT reduces emotional distress resulting in better academic performance.

As NMT technology evolve and improve in quality and accuracy, it can supplement the Malaysian government's Dual Language Programme (DLP) under the "Empower Bahasa Malaysia and Strengthen English" policy.[43] This is particularly useful for teaching Arts stream courses such as linguistics, business, and sociology at various stages of the education system. NMT can help to alleviate the shortage of "language-ready" teachers for DLP classes[44] by allowing instructors to teach in their preferred language while students have the flexibility to listen directly from the instructor or read the projected TTT and STT NMT translation. Furthermore, the deaf and mute community can also benefit from STT NMT, though not in its current form.[45] With the advent of AI, STT NMT can be augmented by incorporating computer-generated gesture avatars to translate spoken to sign language, facilitating communication between non-oral deaf students and their speaking instructors thereby enhancing accessibility and customization of learning.[46] This approach would reduce considerable cost, time, and effort in engaging trained sign language experts, allowing subject matter experts to teach without language barriers. Although sign language translation systems, such as convolutional-based neural network attention module for Malaysian sign language are available,[45] their usage is often limited due to lack of trained corpora and differences in gesture among different sign languages. This shortcoming presents and open-up a new frontier for machine translation researchers to contribute to the wellbeing of the disabled community.

### Limitations and Suggestion for Future Research

This research is constrained by time, resources, and funding, restricting its scope. More importantly, larger samples could improve reliability, consistency, and accuracy of the results. Furthermore, this study confines only to Chinese NNES. Future research should include a more diverse NNES student community to better understand NMT's applications and limitations. Comparing results from various NNES groups would further enrich the insights into NMT's effectiveness.

## CONCLUSIONS

In summary, this study aimed to understand the effect of using NMT in enhancing learner's learning environment in the aspect of enhancing learner-instructor and learner-content interaction leading to an elevated academic performance of NNES. While NMT shows promising results in enhancing learning interactions, overcoming its inherent challenges is crucial for it to be successfully integrated into educational settings. The dynamic nature of user expectations necessitates continuous improvement and adaptation to diverse learner needs. Although NMT can be used in different education levels, it should discretionary be used only upon assessing the language ability of students to prevent deleterious effects on host country students and NNES with higher language proficiency. The findings provide empirical evidence of the correlation between NMT and NNES's academic

performance. In turn, pedagogy centric neural machine translator can be morphed into various education settings, such as the nation's dual language programme and incorporate computer-generated gesture avatars for the benefit of deaf and mute community. With its flexibility, the possibility is endless. However, NMT is not yet ready for critical business dealings, public forums, or teaching of STEM courses that require high accuracy, consistency, and reliability in the interpretation of speech content. Nevertheless, with rapid advancement in neural network and deep learning technology, the widespread use of NMT as application-agnostic may not be far-off in the future.

## BIBLIOGRAPHIC REFERENCES

1. Siddik, Md Abu Bakkar, Akher Ali, Sumon Miah, Mahedi Hasan, Minhaz Ahmed, and Tachlima Chowdhury Sunna. "Psychological disorders among college going students: A post Covid-19 insight from Bangladesh." Journal of Affective Disorders Reports. 2024 Jan; 15(1): 100686. https://doi.org/1016/j.jadr.2023.100686

2. Why are more China parents enrolling children in Malaysia, Singapore Schools? 2022. https://www.scmp.com/week-asia/lifestyle-culture/article/3195153/why-are-more-chinese-parents-enrolling-children

3. Siddharta A. Malaysia: International Students in Higher Education 2022. 2023. https://www.statista.com/statistics/866731/international-students-in-malaysia-by-country-of-origin/#:~:text=In%202022%2C%20there%20were%20approximately,in%20private%20higher%20education%20institutions.

4. Mahmud Z, Amat S, Rahman S, Ishak NM. Challenges for international students in Malaysia: Culture, climate and care. Procedia-Social and Behavioral Sciences. 2010; 7:289-293. https://doi.org/10.1016/j.sbspro.2010.10.040

5. Wang Y, Hu M, Ding R, He P. The dynamic relationship between subjective social status and health: evidence from a Chinese cohort study. British Journal of Health Psychology. 2023 Feb; 28(1):1-21. https://doi.org/10.1111/bjhp.12608

6. Yuhan N, Herasymenko Y, Deichakivska O, Solodka A, Kozlov Y. Translation as a linguistic act in the context of artificial intelligence: the impact of technological changes on traditional approaches. Data and Metadata. 2024; 3:429. https://dm.ageditor.ar/index.php/dm/article/view/263

7. Kumar P. Effect of proficiency in English language on academic performance of post graduate management students of Marathwada region (Maharashtra), India. IOSR Journal of Business and Management. 2014 Jan;16(5):2319-7668. https://doi.org/10.9790/487X-16531016

8. Ekoç A. Using translation as a pedagogical tool: A case study in a writing class at preparatory school. RumeliDE Dil ve Edebiyat Araştırmaları Dergisi. 2019 Aug:346-57. https://doi.org/10.29000/rumelide.606209

9. Shadiev R, Huang YM. Improving student academic emotions and learning satisfaction in lectures in a foreign language with speech-enabled language translation technology. Australasian Journal of Educational Technology. 2022 May 18;38(3):202-13. https://doi.org/10.14742/ajet.7428

10. Khurana D, Koli A, Khatter K, Singh S. Natural language processing: state of the art, current trends and challenges. Multimedia tools and applications. 2023 Jan;82(3):3713-44. https://doi.org/10.1007/s11042-022-13428-4

11. Li H. Deep learning for natural language processing: advantages and challenges. National Science Review. 2018 Jan 1;5(1):24-6. https://doi.org/10.1093/nsr/nwx110

12. Microsoft translator accelerates use of neural networks across its offerings. 2020. https://www.microsoft.com/en-us/translator/blog/2017/11/15/microsoft-translator-accelerates-use-of-neural-networks-across-its-offerings/

13. Zhao Y, Thompson E, Laorga F, Ustimenco T. Important uses of AI in translation. 2022. https://www.aitimejournal.com/important-uses-of-ai-in-translation#:~:text=Machine%20Translation%20uses%20AI%2Dpowered,machine%20translation%20software%20was%20developed.

14. Klimova B, Pikhart M, Benites AD, Lehr C, Sanchez-Stockhammer C. Neural machine translation in

foreign language teaching and learning: a systematic review. Education and Information Technologies. 2023 Jan;28(1):663-82. https://doi.org/10.1007/s10639-022-11194-2

15. Alaboud A. The Positive Effect of Translation on Improving Reading Comprehension among Female Arabic Learners of English as Foreign Language. Arab World English Journal. 2022 Jun;13(2):424-36. https://doi.org/10.24093/awej/vol13no2.29

16. Lee SM. The impact of using machine translation on EFL students' writing. Computer assisted language learning. 2020 Mar 3;33(3):157-75. https://doi.org/10.1080/09588221.2018.1553186

17. Muftah M. The effects of web-based language learning on university students' translation proficiency. Journal of Applied Research in Higher Education. 2023 Nov 22;15(5):1420-39. https://doi.org/10.18844/ijire.v9i2.8496

18. Urlaub P, Dessein E. Machine translation and foreign language education. Frontiers in Artificial Intelligence. 2022 Jul 22;5:936111. https://doi.org/10.3389/frai.2022.936111

19. Al-Musawi NM. Strategic use of translation in learning English as a Foreign Language (EFL) among Bahrain university students. Comprehensive Psychology. 2014 Jan 1;3:10-03. https://doi.org/10.2466/10.03.IT.3.4

20. Lee SM. The impact of using machine translation on EFL students' writing. Computer assisted language learning. 2020 Mar 3;33(3):157-75. https://doi.org/10.1080/09588221.2018.1553186

21. Su W, Li D. The effectiveness of translation technology training: a mixed methods study. Humanities and Social Sciences Communications. 2023 Sep 21;10(1):1-2. https://doi.org/10.1057/s41599-023-02066-2

22. Lake VE, Beisly AH. Translation apps: Increasing communication with dual language learners. Early Childhood Education Journal. 2019 Jul 1;47:489-96. https://doi.org/10.1007/s10643-019-00935-7

23. Shahriar A. The Effectiveness of Machine Translation Using "Google Translate" in English Language Learning in Bangladesh. Pedagogy: Journal of English Language Teaching. 2023 ;11(1):75-88. https://doi.org/10.32332/joelt.v11i1.6355

24. Burbat R, Varela Salinas MJ. Google Translate and DeepL: Breaking Taboos in Translator Training. InFirst International Conference on Research in Multilingualism: Innovation and New Challenges 2018. https://doi.org/10.17398/2340-2784.45.243

25. Mundt K, Groves M. A double-edged sword: the merits and the policy implications of Google Translate in higher education. European Journal of Higher Education. 2016 Oct 1;6(4):387-401. https://doi.org/10.1080/21568235.2016.1172248

26. Bowker L. Fit-for-purpose translation. InThe Routledge handbook of translation and technology 2019 Aug 23 (pp. 453-468). Routledge.

27. Hutchins J. Machine translation and computer-based translation tools: What's available and how it's used. A new spectrum of translation studies. 2003 Mar:13-48.Available from: https://www3.uji.es/~aferna/course4/MT%20process.pdf

28. Lin GH, Chien PS. Machine Translation for Academic Purposes. Online Submission. 2009 Dec.

29. Ardila I. Perception of using machine translation in English subject of Islamic education department students. SKETCH JOURNAL: Journal of English Teaching, Literature and Linguistics. 2021 Dec 26;1(1):23-32. https://doi.org/10.51287/cttl20237

30. Edmonds WA, Kennedy TD. An applied guide to research designs: Quantitative, qualitative, and mixed methods. Sage Publications; 2016 Apr 20.

31. Chufama M, Sithole F. The pivotal role of diagnostic, formative and summative assessment in higher education institutions' teaching and student learning. International Journal of Multidisciplinary Research and

Publications. 2021 Jan;4(5):5-15. Available from: https://ijmrap.com/wp-content/uploads/2021/10/IJMRAP-V4N4P107Y21.pdf

32. Hillson R, Alejandre JD, Jacobsen KH, Ansumana R, Bockarie AS, Bangura U, Lamin JM, Stenger DA. Stratified sampling of neighborhood sections for population estimation: A case study of Bo City, Sierra Leone. PLoS One. 2015 Jul 15;10(7):e0132850. https://doi.org/10.1371/journal.pone.0132850

33. Youngcharoen P, Vincent C. Committee translation approach combined with cognitive interviews: A valuable translation method. Pacific Rim International Journal of Nursing Research. 2016 Apr 27;20(2):91-4. Available from: https://he02.tci-thaijo.org/index.php/PRIJNR/article/view/55686

34. English levels. https://www.britishcouncil.my/english-courses/adults/levels

35. Boamah, E. Improving student information search: A metacognitive approach (Chandos information professional series). The Australian Library Journal. 2015 Oct; 64:348-348. https://doi.org/10.1080/00049670.2015.1100261.

36. Sullivan GM, Feinn R. Using effect size—or why the P value is not enough. Journal of graduate medical education. 2012 Sep 1; 4(3):279-82. https://doi.org/10.4300/JGME-D-12-00156.1

37. Klimova B, Pikhart M, Benites AD, Lehr C, Sanchez-Stockhammer C. Neural machine translation in foreign language teaching and learning: a systematic review. Education and Information Technologies. 2023 Jan; 28(1):663-82. https://doi.org/10.1007/s10639-022-11194-2

38. Shadiev R, Wu TT, Sun A, Huang YM. Applications of speech-to-text recognition and computer-aided translation for facilitating cross-cultural learning through a learning activity: issues and their solutions. Educational Technology Research and Development. 2018 Feb; 66:191-214. https://doi.org/10.1007/s11423-017-9556-8

39. Lee SM. Different effects of machine translation on L2 revisions across students' L2 writing proficiency levels. Learning & Technology. 2022 Sep 5; 26(1): 1–21. https://hdl.handle.net/10125/73490

40. Niehues J, Nguyen TS, Cho E, Ha TL, Kilgour K, Müller M, Sperber M, Stüker S, Waibel A. Dynamic Transcription for Low-Latency Speech Translation. InInterspeech. 2016 Sep:2513-2517. https://dx.doi.org/10.21437/Interspeech.2016-154

41. Huang L, Kern ML, Oades LG. Strengthening university student wellbeing: Language and perceptions of Chinese international students. International Journal of Environmental Research and Public Health. 2020 Aug;17(15):5538. https://doi.org/10.3390/ijerph17155538

42. Herzberg F, Mausner B, Snyderman BB. The motivation to work. Transaction publishers; 2011 Dec 31.

43. Dual Language Programme. Kementerian Pendidikan Negeri Pulau Pinang. 2020. https://jpnpp.moe.gov.my/persekolahan/dual-language-program-dlp

44. Lai A, Ibrahim J. Onus on moe to ensure DLP succeeds, say educators. 2023. https://www.thestar.com.my/news/nation/2023/03/28/onus-on-moe-to-ensure-dlp-succeeds-say-educators

45. Khan RU, Khattak H, Wong WS, AlSalman H, Mosleh MA, Mizanur Rahman SM. Intelligent Malaysian Sign Language Translation System Using Convolutional-Based Attention Module with Residual Network. Computational Intelligence and Neuroscience. 2021; 2021(1):9023010. https://doi.org/10.1155/2021/9023010

46. Gómez Cano C. A., Colala Troya A. L. Artificial Intelligence applied to teaching and learning processes. LatIA. 2023; 1:2. Available from: https://latia.ageditor.uy/index.php/latia/article/view/2

**FINANCING**

The research team would like to thank UOW Malaysia KDU for supporting this language translation project under research grant UOWMKDURG/2023/.

## CONFLICT OF INTEREST

The authors affirm that there is no conflict of interest

## AUTHORSHIP CONTRIBUTION

*Conceptualization:* Shee Mun Yong, Ah Kiaw Phan.
*Data curation:* Shee Mun Yong, Ah Kiaw Phan.
*Formal analysis:* Shee Mun Yong, Ah Kiaw Phan.
*Research:* Shee Mun Yong, Ah Kiaw Phan.
*Methodology:* Shee Mun Yong, Ah Kiaw Phan.
*Project management:* Shee Mun Yong, Ah Kiaw Phan.
*Resources:* Shee Mun Yong, Ah Kiaw Phan.
*Drafting - original draft:* Shee Mun Yong, Ah Kiaw Phan.
*Writing - proofreading and editing:* Shee Mun Yong, Ah Kiaw Phan.

https://doi.org/10.56294/saludcyt20251426