

Developing and Validating a Scale for Engagement in Writing with Machine Translation

Mariko Yuasa*

Graduate School of Foreign Language Education and Research,
Kansai University, Osaka, Japan

Abstract

Machine translation (MT) is increasingly used by second language writers worldwide, but its use by low-proficiency English as a foreign language (EFL) students has been criticised for causing cognitive disengagement in writing. Moreover, research on MT's impact in this context is limited. As students' engagement in the writing process is essential for using MT for language learning, this study developed and validated a scale with multidimensional engagement constructs to understand how low-proficiency EFL students engage in writing using MT. A 24-item instrument covering behavioural, cognitive, affective, social, and agentic subscales was administered to 773 Japanese university students, mostly at the Common European Framework of Reference A2 level. Exploratory factor analysis with half the participants identified five engagement constructs, with cognitive engagement subdivided into pre-editing and post-editing, with agentic engagement being excluded because of low factor loadings. This hypothesised five-factor model was tested with the remaining participants using confirmatory factor analysis, which yielded satisfactory reliability coefficients, as well as construct, convergent, and discriminant validity. Concurrent validity was weakly supported, with a near moderate correlation between engagement and self-efficacy. The scale offers insight into Japanese students' engagement with MT and provides educators with fresh insights into their use of MT.

Keywords: Engagement, Machine translation, Second language writing, Self-efficacy, Exploratory factor analysis, Confirmatory factor analysis

*Corresponding Author E-mail: maritama0617@gmail.com

Background to the Study

Recent studies have increasingly documented the positive impact of machine translation (MT) on second language (L2) writing in higher education (Lee, 2023), particularly for intermediate or advanced learners (e.g., Klimova et al., 2022). In Japan, many university students are at the Common European Framework of Reference (CEFR) A2 level, raising questions about the justification for using MT. As “MT-enabled L2 writing is becoming a norm” (Jiang et al., 2024, p. 1), teachers play a crucial role in helping low-proficiency learners refine their MT strategies (Kern, 2024). To create such learning opportunities, we must first understand how learners engage with MT. Applying the multidimensional constructs of engagement to quantitative surveys could provide an overview of how low-proficiency English as a foreign language (EFL) students engage in writing with MT and address the mismatch between the students’ needs and current classroom practices. Engagement broadly refers to active participation in learning, encompassing behavioural, cognitive, affective, and social factors (Hiver et al., 2024). Motivated by Zhou et al.’s (2021) call to extend engagement research to the potential role of technology in facilitating L2 learning, this study developed a scale called Engagement in Writing with MT to explore the following research questions:

RQ1. What is the factor structure of the Engagement in Writing with MT scale?

RQ2. How reliable and valid is the Engagement in Writing with MT scale?

Method

Participants

The participants included 773 (males = 387; females = 372) Japanese EFL students recruited from a private university through convenience sampling. They were non-English majors, mostly at CEFR A2, with 392 freshmen and 381 sophomores.

Instrument

An engagement scale was developed for this study. The Self-efficacy for Writing with MT scale (Yuasa & Takeuchi, 2022) was used to test the concurrent validity of the developed scale, based on the theory that higher self-efficacy correlates with greater engagement in learning tasks (Linnenbrink & Pintrich, 2003; Schunk & DiBenedetto, 2022).

Procedures

Five engagement constructs for L2 writing were defined: behavioural (the time and effort spent using MT), cognitive (mental effort and strategic use to improve the output of MT), affective (attitudes and positive feelings towards MT), social (interaction with peers and others to use MT better), and agentic (conscious and proactive personalisation of MT use). A 24-item questionnaire using a 5-point Likert scale was developed and administered to the 773 participants via Google Forms.

Data cleaning reduced the sample to 708 participants, which were then split into two groups. Following checks for factorability and sampling adequacy, an exploratory factor analysis (EFA) with maximum likelihood (ML) was conducted on half the participants ($n = 354$) using JASP software (version 0.18) to identify the factor structure. A confirmatory factor analysis (CFA) was then conducted with the other half ($n = 354$) to test the latent constructs identified in the EFA. Internal consistency as well as the construct, convergent, discriminant, and concurrent validity were examined.

Results

Factor Structure

Kaiser's criterion, a scree plot, and parallel analysis confirmed a five-factor solution for the engagement scale: behavioural (e.g., "I rely heavily on MT when writing in English"), affective (e.g., "Using MT makes writing in English less painful"), social (e.g., "I ask my friends, teachers, or someone outside school how to use MT"), and cognitive, which was subdivided into pre-editing (e.g., "I use easy to understand Japanese in the source text") and post-editing (e.g., "I check the English in the MT's output and self-correct errors if necessary"). Agentic engagement was not extracted because of low factor loadings (>0.50) (Hair et al., 2019).

Reliability and Validity

Table 1 shows the factor structure, reliability coefficients, and average variance extracted (AVE). McDonald's omega and Cronbach's alpha coefficients met the 0.70 threshold, confirming the internal consistency. Construct validity was established with good indices of model fit: $\chi^2 = 265.34$ ($df = 125$, $p < 0.01$), root mean square error of approximation = 0.06 (90% confidence interval (CI): 0.05–0.07) (<0.08), comparative fit index = 0.95 ($>.90$), Tucker–Lewis index = 0.94 (>0.90), and standardised root mean square residual = 0.05 (<0.08). Convergent validity, as measured by AVE, was acceptable (>0.50)

Table 1. Factor structure, reliability, and AVE of engagement in writing with MT

Engagement Factors	Items	Factor Loadings	Residual Variances	McDonald's ω	Cronbach's α	AVE
Behavioural	4	0.88	0.23	0.90	0.89	0.64
	2	0.85	0.27			
	5	0.84	0.29			
	3	0.77	0.41			
	1	0.60	0.65			
Cognitive (pre-editing)	16	0.79	0.38	0.73	0.73	0.41
	15	0.76	0.43			
	17	0.53	0.73			
	19	0.51	0.74			
Cognitive (post-editing)	6	0.84	0.30	0.86	0.86	0.68
	7	0.82	0.32			
	8	0.81	0.35			
Affective	22	0.81	0.34	0.76	0.75	0.51
	21	0.78	0.39			
	23	0.56	0.69			
Social	11	0.80	0.37	0.78	0.77	0.55
	10	0.76	0.43			
	13	0.63	0.61			

Note. AVE = average variance extracted.

(Hair et al., 2019). Discriminant validity was confirmed when the square root of the AVE for each construct exceeded the inter-construct correlations. Concurrent validity was weakly supported, as the correlation between the engagement and self-efficacy scales was near moderate ($r = 0.36$; $p < 0.001$).

Discussion

EFA identified a five-factor structure for the engagement scale (RQ1), consistent with the proposed theoretical construct. However, agentic engagement was not observed, possibly because Japanese university teachers generally do not support students' use of MT (Yamada et al., 2021), which may inhibit students' proactive or agentic behaviour (Reeve & Shin, 2020). This could explain why this factor was not found. Agentic engagement is “in its infancy” (Patall, 2024, p. 7), requiring further research to validate this construct (Sinatra et al., 2015).

Validation of the scale (RQ2) showed satisfactory reliability and validity. Weak support for concurrent validity, resulting from an almost moderate correlation between self-efficacy and cognitive engagement, might be caused by the students' perceptions of the performance of MT; some may believe MT is highly accurate and will therefore use few error correction strategies. This aligns with Koltovskaia's (2020) finding that over-reliance on Grammarly's automated feedback led to limited use of strategies and vice versa.

Conclusion

This study is the first to use a quantitative survey method to extend the engagement framework and explore how low-proficiency EFL students engage in writing with MT. The context-specific scale can help educators understand the use of MT by these students and guide the creation of learning opportunities through integrating MT into EFL classrooms, better addressing the needs of artificial intelligence-assisted language learning.

Acknowledgements

This work is supported by the 36th Research Grant from Eiken Foundation of Japan.

References

- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2024). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research*, 28(1), 201–230. <https://doi.org/10.1177/13621688211001289>
- Jiang, L., Yu, R., & Zhao, Y. (2024). Theoretical perspectives and factors influencing machine translation use in L2 writing: A scoping review. *Journal of Second Language Writing*, 64, 101099. <https://doi.org/10.1016/j.jslw.2024.101099>
- Kern, R. (2024). Twenty-first century technologies and language education: Charting a path forward. *The Modern Language Journal*, 108(2), 515–533. <https://doi.org/10.1111/modl.12924>
- Klimova, B., Pikhart, M., Benites, A. D., Lehr, C., & Sanchez-Stockhammer, C. (2022). Neural machine translation in foreign language teaching and learning: A Systematic review. *Education and Information Technologies*, 28(1), 663–682. <https://doi.org/10.1007/s10639-022-11194-2>

- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, 44, 100450. <https://doi.org/10.1016/j.asw.2020.100450>
- Lee, S.-M. (2023). The effectiveness of machine translation in foreign language education: A systematic review and meta-analysis. *Computer Assisted Language Learning*, 36(1–2), 103–125. <https://doi.org/10.1080/09588221.2021.1901745>
- Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading & Writing Quarterly*, 19(2), 119–137. <https://doi.org/10.1080/10573560308223>
- Patall, E. A. (2024). Agentic engagement: Transcending passive motivation. *Motivation Science*, 10(3), 222–233. <https://doi.org/10.1037/mot0000332>
- Reeve, J., & Shin, S. H. (2020). How teachers can support students' agentic engagement. *Theory Into Practice*, 59(2), 150–161. <https://doi.org/10.1080/00405841.2019.1702451>
- Schunk, D. H., & DiBenedetto, M. K. (2022). Self-efficacy and engaged learners. In A. L. Reschly, & S. Christenson (Eds.), *Handbook of research on student engagement* (2nd ed., pp. 155–170). Springer.
- Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, 50(1), 1–13. <https://doi.org/10.1080/00461520.2014.1002924>
- Yamada, M., Langliz, H., Oda, T., Morita, T., Tamura, H., Hiraoka, Y., & Irie, T. (2021). A preliminary survey on the use of machine translation in English education at Japanese universities. *Invitation to Interpreting and Translation Studies*, 23, 139–156. http://honyakukennyu.sakura.ne.jp/shotai_vol23/No_23_007-Yamada_et-al.pdf
- Yuasa, M., & Takeuchi, O. (2022). *Strategic use of Google Translate by Japanese university students of English* [Conference presentation]. *Situating Strategy Use: Strategic Learning in an Uncertain world (SSU4)*, Wellington, New Zealand.
- Zhou, S. A., Hiver, P., & Al-Hoorie, A. H. (2021). Measuring L2 engagement: A review of issues and applications. In P. Hiver, A. H. Al-Hoorie, & S. Mercer (Eds.), *Student engagement in the language classroom* (pp. 75–98). Multilingual Matters.