Automated Scoring of EFL Learners' Written Performance: A Torture or a Blessing?

[1]Roya Khoii, [2]Amir Doroudian

[1]Islamic Azad University, North Tehran Branch, [2]Islamic Azad University, South Tehran Branch (Iran)
roya_kh@yahoo.com, amirdoroudian@aol.com

Abstract

Technology has been a game changer in the quest to improve language teaching and learning, and with the advancement of Artificial Intelligence (AI), this influence has become even more decisive. In line with this enterprise, different types of Automated Writing Evaluation (AWE) software, which score essays and generate feedback using AI, have been developed to meet the challenge of evaluating learners' written performance. AWE programs provide learners with immediate scores and diagnostic feedback on their essays, as well as convenient writing tools to facilitate the writing process. The validity and reliability of the scores such programs generate, along with the scoring engines behind them have been investigated in an array of studies. However, research on using AWE in EFL classrooms and its outcomes is sparse in the literature. Moreover, studies are not conclusive as to the effect of using such programs on improving L2 learners' writing skill. Finally, with the growing availability of commercial AWE products in recent years and their subsequent educational impacts, the question of their usefulness as an instructional tool has become ever more relevant. Thus, the present research is conducted to compare the effects of automated and human scoring of written performance in a second language with each other. The participants of this study consist of 30 foreign language learners equally divided in two essay writing classes. The experimental class uses the AWE program My Access! ® as the scorer of essays and students' source of feedback, while the control class uses a teacher for similar purposes. Both classes are taught by the same teacher using the same teaching materials and the same criteria for scoring the essays. The 10-week treatment period commenced with an essay writing pretest to measure the participants' writing ability at the outset of the study. At the end of the treatment, they will receive an essay writing posttest in order to evaluate their level of progress in terms of writing. A questionnaire will also be given to the experimental participants to learn about their experience regarding the use of AWE in their class. The required statistical analyses will be performed following the completion of data collection, and the results will be known in due course.