Chapter 9

# Human vs. AI:
## The Use of ChatGPT in Writing Assessment

**Betul Kınık**
*Inonu University, Turkey*

**Hande Çetin**
https://orcid.org/0000-0002-2710-4083
*Tokat Gaziosmanpasa University, Turkey*

## ABSTRACT

*The current study seeks to investigate whether ChatGPT 3.5 can be used as an aid to help diminish the teachers' workload in assessing writing. To this aim, a mixed-methods research design was employed for the study. Randomly selected, 20 descriptive essays written by freshman student teachers of English Language Teaching were scored by an experienced human rater and ChatGPT 3.5. An adapted 'descriptive essay rubric' by the researchers was used to assess the descriptive essays of the student teachers. The quantitative aspect of the study involved frequency and percentage analysis, while the qualitative dimension centered on analyzing the written feedback provided by both ChatGPT and the human rater. The findings showed that there is a disagreement between ChatGPT 3.5 and the human rater. Furthermore, there are some problems with the written feedback it provides. It is clear that it is rapid in terms of providing feedback. Thus, it is recommended that ChatGPT 3.5 can be employed as a tool under the supervision of teachers.*

## INTRODUCTION

The literature consistently emphasizes the importance of students practicing writing extensively to enhance their ability to express themselves effectively. However, it is acknowledged that teachers often face time constraints, hindering their ability to provide timely feedback and evaluate student assignments. This challenge has led to the expansion of the Automated Writing Evaluation (AWE), also accentuated as Automated Essay Evaluation and Automated Essay Scoring (Huawei & Aryadoust, 2023) and its supporters. AWE leverages Artificial Intelligence (AI) technology to rapidly score written work (Cushing Weigle, 2010; Warschauer & Grimes, 2008). "The sheer number of hours commenting on student papers is reduced dramatically when instructors can rely on automated electronic feedback systems" (Ware & Warschauer, 2006, p. 108). Whereas the studies have shown that AWE can be used as an aid for teachers to score writing (Wilson & Roscoe, 2020), it is worth noting that there remain a number of concerns raised, especially among writing educators (Cushing Weigle, 2010).

AWE systems make use of AI (Steiss et al., 20223), which can be specified as "natural language processing (NLP) and machine learning techniques" (Uto, 2021), examine written texts, and promptly produce ratings based on the writing quality. Furthermore, they provide written feedback to improve overall and specific aspects of writing (Cushing Weigle, 2010). While the current functionalities of AWE systems are impressive, it's worth tracing their evolution over the decades.

It was 1960s when Page (1966) mentioned the possibility of scoring the essays in an automated way with their project called 'Project Essay Grade'. Three decades later, with the developments in computer and the Internet created possibilities for marketing globally (Warschauer & Grimes, 2008). Education Testing Service (ETS) developed a system which was called *e-rater* to score the essays written in TOEFL IBT (Cushing Weigle, 2010). Simultaneously, Intellimetric was developed by Vantage Learning. Another engine to score high-stake writing tasks was Intelligent Essay Assessor which was developed by a group of academics used a technique called latent semantic analysis. Pearson, a publishing company, later acquired this technology for their tests (Warschauer & Grimes, 2008). Apart from these, there are AWE tools which were designed for classroom use, including Writer's Workbench, MY Access!, WriteToLearn, Criterion, and RWT (Saricaoglu, 2015). AWE systems make use of natural language processing (NLP) (Wison & Roscoe, 2020).

AWE tools have been investigated from different aspects, such as the validity of the scoring in high-stakes tests (Attali & Burstein, 2006; Cushing Weigle, 2010); effectiveness of AWE (Stevenson & Phakiti, 2014; Wilson & Roscoe, 2020); AWE's reliability (Shermis, 2014). Furthermore, the researchers have started to examine how Chat Generative Pre-trained Transformer (ChatGPT) can be used for assessment

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/human-vs-ai/336060

## Related Content

Professional Support Networks in TESOL: Sociolinguistic Approaches to Professional Development in ESOL Teacher Education
Bahar Otcu-Grillmanand JungKang Miller (2022). *Research Anthology on Applied Linguistics and Language Practices (pp. 768-784).*
www.irma-international.org/chapter/professional-support-networks-in-tesol/303048

Paratext Analysis of Patronage Factors: An Exploration of Howard Goldblatt's Translation of Mo Yan's Life and Death Are Wearing Me Out
Hu Liu (2021). *International Journal of Translation, Interpretation, and Applied Linguistics (pp. 30-44).*
www.irma-international.org/article/paratext-analysis-of-patronage-factors/266314

Preparing Preservice Teacher Candidates for Educating Culturally and Linguistically Diverse Learners
Alpana Bhattacharya (2020). *Beyond Language Learning Instruction: Transformative Supports for Emergent Bilinguals and Educators (pp. 76-103).*
www.irma-international.org/chapter/preparing-preservice-teacher-candidates-for-educating-culturally-and-linguistically-diverse-learners/241662

A Critical Perspective on the Challenges for Blended Learning and Teaching in Africa's Higher Education
Alfred T. Kisubi (2014). *Cross-Cultural Interaction: Concepts, Methodologies, Tools, and Applications (pp. 1043-1066).*
www.irma-international.org/chapter/a-critical-perspective-on-the-challenges-for-blended-learning-and-teaching-in-africas-higher-education/97518

Bilingual/Bicultural Paraprofessionals as Brokers of English Learners' Family-Teacher Relationships
Eva I. Díaz, Diana Gonzales Worthen, Conra D. Gistand Christine Smart (2021). *Bridging Family-Teacher Relationships for ELL and Immigrant Students (pp. 46-64).*
www.irma-international.org/chapter/bilingualbicultural-paraprofessionals-as-brokers-of-english-learners-family-teacher-relationships/266919