

Title	Evaluating the Quality of ChatGPT 3.5-Generated Academic Essays
Author(s)	Bui, Minh Ngoc
Citation	
Issue Date	2024-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/19349
Rights	
Description	Supervisor: Kim Eunyoung, 先端科学技術研究科, 修士(知識科学)

Evaluating the Quality of ChatGPT 3.5-Generated Academic Essays

2210423 BUI, Minh Ngoc

This research investigates the writing quality of academic essays produced by ChatGPT and the factors influencing their effectiveness. Despite the growing use of AI-generated content in education, there is a scarcity of studies on its academic writing proficiency. The objective of this study is to address this gap by evaluating different facets of essay quality, including grammar, coherence, originality of ideas, content development, plagiarism detection, and citation precision.

This study combines a thorough analysis of AI-generated essays with statistical evaluations to investigate ChatGPT's writing abilities. By exploring a diverse range of essay topics and evaluating various writing components, this research offers detailed insight into ChatGPT's proficiency in academic writing. The outcomes hold considerable significance for educators, students, education policymakers, and AI developers.

The primary research goal is to assess the quality of ChatGPT-generated essays across different themes and writing assessment criteria. Sub-objectives include evaluating grammar, coherence, originality of ideas, content development, plagiarism prevention, and reference accuracy in AI-generated essays, as well as identifying factors influencing ChatGPT's writing quality.

The study methodology consists of four main steps. Initially, *Data Generation* involved creating a variety of essays using ChatGPT on diverse academic topics. Subsequently, *Scale Development* entailed constructing a comprehensive rubric to assess essay quality based on criteria such as grammar, coherence, originality, content development, plagiarism prevention, and reference accuracy. During the *Data Evaluation* phase, each essay was reviewed using this rubric to ensure consistent and reliable assessments. Finally, *Data Analysis* employed statistical methods, including one-way ANOVA and correlation analysis, to examine variations in writing quality across different themes and identify significant correlations among the evaluation criteria.

The study highlights ChatGPT's strong performance in grammar and coherence, which is particularly beneficial for non-native English speakers. However, shortcomings are observed in content development and personalized conclusions. The research identifies critical thinking abilities and writing proficiency as key components within the writing quality assessment framework, showing a positive correlation with ChatGPT's writing quality. Additionally, consistent writing competence by ChatGPT across various subjects is noted, indicating its effectiveness regardless of the topic.

The findings of the study offer significant insights for educators seeking to seamlessly incorporate AI tools into the curriculum to enhance student learning. Educational policymakers within schools can utilize these results to establish protocols ensuring the ethical and effective use of AI in educational settings. Furthermore, AI developers can capitalize on the identified shortcomings to improve the functionality and performance of forthcoming AI writing tools.