



Available Online

# Journal of Education and Social Studies

ISSN: 2789-8075 (Online), 2789-8067 (Print)

<https://www.scienceimpactpub.com/jess>

## ASSESSING THE EFFECTIVENESS OF CHAT GPT IN ENGLISH LANGUAGE AND LITERATURE: AN EXPERIMENTAL INVESTIGATION INTO AI-GENERATED RESEARCH CONTENT

Bushra Munawar<sup>1</sup>, Hina Rafique<sup>2,\*</sup>, Nida Mushtaq Khan<sup>3</sup>

<sup>1</sup> Southern Hampshire University, USA

<sup>2</sup> Govt Graduate College For women, Satellite Town, Gujranwala, Pakistan

<sup>3</sup> Department of Learning Sciences, The University of Oklahoma, Norman, Oklahoma, USA

### ABSTRACT

ChatGPT is an artificial intelligence model, highly proficient in English language and literature. This research explored the functionality of ChatGPT in producing research-oriented texts of the English language and literature in higher education institutions. It consisted of careful analysis of the quality, coherence, and relevance of ChatGPT-produced content in terms of its capability to develop research topics, abstracts, research methodologies and theoretical-framework-based data analysis, and conclusions. Furthermore, the study focused on its role in academic context by drawing upon a comprehensive analysis of the strengths and limitations of ChatGPT in the research domain. Hence, the analysis consisted of scrutinizing the data for completeness, accuracy, and relevance of the content produced by ChatGPT administering the CRAAP Test as the designed measure of performance of ChatGPT. Results showed a high proficiency of ChatGPT in originating research ideas and data explication but limiting response in composition of comprehensive abstract writing, research methodologies and principle design, especially when implementing practical implementation of overall research components. Furthermore, the study proposed that researchers should approach AI technology with a critical mindset. Contextually, the incorporation of human linguistic and literary expertise is recommended in utilization of ChatGPT in research studies and academic term papers. These measures will help maintain the quality and integrity of research and academic studies.

**Keywords:** ChatGPT; Artificial Intelligence; Language model; Research-content-generation; Scholarly inquiry; Research writing; Human oversight.

\* Email: [hina.gcw@gmail.com](mailto:hina.gcw@gmail.com)

© The Author(s) 2024.

<https://doi.org/10.52223/jess.2024.5325>

Received: June 16, 2024; Revised: September 14, 2024; Accepted: September 21, 2024

This is an open-access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### INTRODUCTION

The innovations in AI (artificial intelligence) have revolutionized various fields, such as media and communication studies, translanguaging and language teaching, and language generation and academic research. Over the last few years, AI-based natural language processing models have captured the attention of many scholars, especially because of their capability to produce text akin to human-like text in many different domains. However, among these approaches, ChatGPT is the major performer and it can produce grammatically correct, and contextually relevant answers using diverse learning techniques. Recently, the researchers and graduate and post graduate level students have been observed doing their research projects, and writing term papers with the help of ChatGPT. This phenomenon seems to be in vogue in higher education institutes; public and private graduate and post graduate colleges and universities. Contextually, the prime purpose of this research is to understand the effectiveness of ChatGPT as an educational tool for writing research studies and academic papers in English language. This study aims to

answer the critical questions about the usefulness of ChatGPT for academic research by examining its efficiency through experimental investigation of selected texts.

Two primary objectives guide this research: firstly, to gauge the effectiveness of ChatGPT as an AI model in generating research-oriented materials within English language and literature domain; and secondly to critically evaluate the content produced by ChatGPT, focusing on aspects such as quality, coherence, and relevance. These objectives align with the goal of this study: to identify the strengths and limitations of AI-driven tools in supporting scholarly processes. Considering the increasing reliance on AI as a research assistant, it is crucial to assess these tools' performance in meeting academic rigor standards (Rafique et al., 2024). In recent times, ChatGPT, a transformative language-generation model, has emerged as a powerful tool with applications that extend beyond many fields. Trained on vast amounts of text data, it demonstrates an exceptional ability to interpret and produce text in natural language. Its capacity to generate well-structured and relevant sentences makes it a valuable asset for copywriting, summarization, and dialogue generation (Mailjile et al., 2024). Yet, the question remains: to what extent can it be considered a research-oriented data source, capable of analysing the language of social sciences and literature.

Digitalization has brought about an easy access to academic resources and the expansion of online platforms, dramatically altering scholarly communication. The researchers have plenty of information, instruments, and various methods. AI language models such as ChatGPT can greatly enhance research efficiency by simplifying the process of examining literature and creating content. These models are particularly instrumental useful in summarizing data, generating hypotheses, and producing research papers. Researchers can save time and resources by utilizing machine learning technology and linguistic processing while producing high-quality work (Kolade et al., 2024). While the use of AI in academic research is a powerful tool, it is important to acknowledge the concerns raised about ethics, and quality and credibility of the data. The question of whether digital mechanisms can ever replicate the quality and reliability of human-generated content is a valid one. If translation quality and fluency are the main factors, the performance of AI models like ChatGPT must be rigorously examined (Khlaif et al., 2023). This assessment involves several criteria, including their understanding of abstract ideas, ability to persuade using sound reasoning and logical arguments, and knowledge of research skills like research through quotations etc. By addressing these concerns, we can ensure the integrity of AI-driven research.

An experimental study of "ChatGPT efficiency" is a formal strategy for identifying whether this AI tool can produce content for research purposes. Researchers can use factors from the experiment, including relevancy and style consistency, to reinforce a survey supported by ChatGPT content that mimics human-authored content in terms of accuracy and coherence. Similarly, experts and functioning users can evaluate and enhance the functionality of ChatGPT by offering worthwhile comments on the real-life uses of this technology in research (Song & Song, 2023). This study proposes the place of AI in education in English language and literature by investigating the effectiveness of the tool ChatGPT in crafting research-oriented content for research studies; research articles, case studies and critical analyses etc. It intends to collect and analyse data results to assess ChatGPT through an empirical trial and evaluation. The pivotal concern is to list ChatGPT's strengths and limitations. The research outcome will present the utility of ChatGPT as a tool for scholarly communication. In this context, this study aims to justify these questions and promote the future of AI-driven research assistance and participation in discussions concerning tech and science (Young & Shishido, 2023).

### **Research Statement**

While conducting research, the researchers have to write abstracts, literature reviews, a good introduction and theoretical frameworks. They also have to analyse textual data in context of selected theoretical insights. This research aims to accomplish two main objectives: the first will explain how proficient ChatGPT is as a language model, and the second will highlight the quality, coherence, and relevancy of the

content it generates. This evaluation aims to demonstrate the advantages and disadvantages of using ChatGPT as an educational tool using a more structured framework and a critical analysis. Besides, the objective is to maintain that artificial intelligence technology can contribute to English language and literature research only when its design is well-developed.

### **Research Objectives**

The research objectives of this study are the following:

1. To evaluate the effectiveness of ChatGPT in generating research-based content within the realm of English language and literature
2. To ensure the quality, coherence, and relevance of ChatGPT-generated content for research purposes.

### **Significance of the Study**

This research focuses on utilizing AI language models to conduct research in English language and literature domains. Through examination of the research- focused writing of ChatGPT- generated-content, the quality, comprehension, and relevance of the content will be evaluated. This evaluation can guide researchers about the usefulness of AI language models in language learning and literature research. Research into this area can provide some useful insights into the efficiency of AI-powered tools in writing research-based assignments in higher education institutions at graduate and post graduate levels. This study will also provide valuable data regarding the role of the language modal of the artificial intelligence in conducting literature searches, aiding content generation, and facilitating scholarly exchange of ideas. It will also provide the direction for future utilization of this tool in research and education field.

### **Delimitations**

This research is limited by its scope and methodology, and these factors are the main drawbacks that should be observed. Primarily, the study is limited to analysing the productivity of the AI tool ChatGPT concerning both the qualitative and quantitative academic content provided within the context of the English language and literature. ChatGPT is versatile and can be employed in multiple domains, so this research is limited to this area. Furthermore, the research approach depends on qualitative research methods, utilizing ChatGPT, creating content for the agreed research topics, and then analysing the outputs against the CRAAP test. Such a methodological approach enforces the study to provide a qualitative qualitative evaluation of ChatGPT outputs that are evaluated with predetermined measures that may narrow the scope of the research to the applicability of the selected impartiality method. Moreover, using the CRAAP Test as the sole measure of performance of ChatGPT may provoke a deficiency in the assessment because it only focuses on other aspects of the performance of ChatGPT, such as language nuances or style appropriateness.

### **LITERATURE REVIEW**

AI has attracted scholars because of its use in language generation and academic research. The literature explains the scope and challenges of AI-driven tools. OpenAI builds ChatGPT, one of the most advanced language models today. It is based on the transformer architecture that makes generating human-like text responses to the given prompts possible. Since then, ChatGPT has gained great recognition among researchers across different disciplines because of its advanced adaptability to context, sustaining cohesion, and producing fluent texts. Thanks to its quick generalization, Reed (2023) showed that ChatGPT can be applied to multiple tasks with minimal fine-tuning.

One promising application of AI language models such as ChatGPT within academia is improving different aspects of the scholarly communication process. Niloy et al. (2024) proved the capability of language models to learn multitasks without supervision, including translation, text summarization, and question answering. By refining the pre-trained models for the specific domain or task, scientists can effectively use AI language models produced according to their research needs (Nazeer et al., 2024).

On the other hand, AI models can only be trusted solely for producing scientific research within the English language and literature niche with further scrutiny. Laing et al. (2024) discussed what is most effective in improving the interactions between language learners and machines via domain-related fine-tuning. Although generic language models like ChatGPT are good for general text generation purposes, their accuracy may fluctuate when used in specialized areas that have their linguistic characters and conventions.

Chaka (2023) pinpoints that AI-generated content must be coherent and relevant in academic research. The authors suggested a neural topic modelling approach built on document clustering to enhance the coherence and relevancy of generated texts. Using semantic similarity, AI language models can tailor their content to themes of interest and predict future trends. As an example, the research of Mindner et al. (2023) clearly shows that the contextual competency of recommendation systems is the main factor in effective scholarly communication. Using hierarchical compound attention networks, one can create AI systems that cite referred literature and sources directly related to the study's main subject. Combining AI recommendation systems and ChatGPT will result in the information being valuable and at a deeper level (Haleem et al., 2022). Talking about the pragmatic competence of ChatGPT, the researchers Nazeer et al., argue about their research, "through a systematic assessment of the capability of ChatGPT to determine the context, recognize tone, and effectively process complex linguistic aspects like irony, metaphor, and indirect requests; this study makes meaningful contributions to the domain of AI and NLP (2024, p 426). They further opine, "Nevertheless, its implications go further for developers of AI language models as well as researchers in line with and end-user" (Nazeer et al., 2004).

While the researchers are evaluating the accuracy and logic of AI-generated counter searchers, questions raised about ethics on AI applications in academic research is another issue being investigated? As Imran and Shambarrouf (2023) highlighted, there is a threat of AI language models that may echo the biases in data they have learned. The scholars recommended using more open and transparent models of AI advancement to avoid the chance of biased or erroneous data reinforcement.

In addition, the impact of scholars' writing and attribution errors has been similarly studied using studies conducted in the US. Herbold and colleagues (2023) have recently explored whether of academic authors would go against using AI tools that automatically generate text and the implications for authorship. The research highlights that the writers either welcome them as a tool or they, on the other hand, raise authorship problems and originality issues (Wang et al., 2023). AI language models and ChatGPT are especially persuasive within the literature in education that administer AI the autonomy to be used in all research projects, such as content generation, literature reviews, and citations. However, mining would be impacted by problems like lexical gaps, cohesiveness, relevance, and ethical dilemmas, which require further studies. Overcoming these issues using scientific approaches such as experiments and analyses is the only tool that will help upgrade the knowledge base of AI in science, and it is recommended to create more effective AI-based academic communication devices (Tirado-Olivares et al., 2023).

## **METHODOLOGY**

This qualitative study focused on the functionality of the ChatGPT AI tool and its power to generate qualitative and quantitative research content. The researchers devised two research topics: The textbook comprises two sections: "textual analysis of the poem The Stormy Sky, Written by Imran Nazeer" and "Stylized analysis of a short story The Winepress, written by Josef Essberger" (Nazeer, 2022; Essberger, 2021). Being provided with various possible prompts, ChatGPT was directed to create content. ChatGPT was also asked to think up a research topic for English semantics and make a research proposal on the suggested topic. Later on, the results were then analysed, using the CRAAP Test (2004), a method developed by Librarian Sarah Blakeslee, that determines how fresh the source is, how relevant it is, and if it has an author with authority, credible information, and reliable purpose. The investigation took this

measurement as its basis, and thus, the analysts wanted to determine the quality and truthfulness of the outputs of the system in context of the research content.

Those research findings showed how ChatGPT could work to create content that links with research. This goal was achieved by conducting both qualitative and quantitative output analysis. The effectiveness of ChatGPT in addressing particular research topics was evaluated through the analysis of these outputs. Using the CRAAP Test, the researcher evaluated the currency, relevance, authority, accuracy, and purpose of the produced content, drawing inferences about its credibility and suitability for academic analysis (Esprarago-Kalidas. 2021).

## **DATA ANALYSIS AND RESULTS**

The researcher assigned different tasks to ChatGPT to complete them including “suggestion of a research topic”, “writing a research proposal”, and “to conduct the research on the given research topic.” ChatGPT completed all the prompts but there were some issues with the content generated by ChatGPT. The detailed analysis is given below:

### **ChatGPT-Generated Content Analysis**

#### ***Research Topic***

The researchers interacted with ChatGPT by asking for some topics that could fall under the field of Semantics. ChatGPT proposed a compelling subject. This exchange shows the fact that it emphasizes the possibility of making topics more refined. Offering other prompts to ChatGPT in the divergent areas shows that it is “smart” and can handle input smartly. Nevertheless, when given the specific task of creating research topics from the existing literature, ChatGPT displayed its expert skills by giving equally impressive alternatives. Such collaborations render more evidence of its functionality to produce purposeful research topics, thus demonstrating how it is a versatile tool with which scholars can investigate and inquire.

#### ***Abstract***

Abstracts prepared by ChatGPT lack fundamental elements critical for the research studies. Interestingly, they also do not contain elements referred to as research methodology, data analysis, and findings, which are, more or less, the core of a research abstract. Thus, the failure of ChatGPT to condense the whole text into brief abstracts shows its restricted function, as it cannot go beyond the main points. Abstracts function like a précis, providing readers with essential details concerning the nature of the research, the methods applied, the findings obtained, and the significance of the results. The abstracts become less comprehensive and well-rounded in their overall impression if vital components are missing, implying that the system cannot produce detailed; rich and informative abstractions.

#### ***Introduction***

The lack of in-text citations mentioned in the introductions of the ChatGPT-generated papers for the various research articles signifies the inability of AI to produce truly academic content. Despite ability of ChatGPT to create captivating and comprehensive introductions, the content could be more relevant to academia due to the need for proper citations. Citations within academic writing perform the fundamental function of attribution and serve to attribute the platform from which the text's information and ideas are sourced. Quoting without citations, the introduction does not meet the standards of genuine research and may be considered by others to be lacking scholarly authority. Thus, ChatGPT presents extensive skills in writing introductions for research studies, when provided the tight clues. However, a lack of appropriate citations occasionally results in the production of unacceptable and untrustworthy content.

#### ***Research Objectives***

The type of research objectives furnished by ChatGPT reveals the limitation that it cannot give out detailed and focused research goals. Well-defined and precise research goals are key elements of a good research

design. They are meant to ensure that the area of interest is focused and specific and that, thus, the research project addresses particular questions or hypotheses. The specificity of an objective plays a role in bringing, and this can either be a success factor or a disaster. Consequently, the ability of ChatGPT to develop research content is limited by providing generic objectives, and therefore, human intervention is crucial to reconcile the clarity for effective research.

### ***Literature Review***

The replication of content in the introduction and literature review and the inability of ChatGPT to put together and cite the previous literature on the topic are signs of the limitation of its access to literature and integration of the previous literature. A good literature review encompasses a thorough synthesis of the existing research on the topic; for instance, various studies involve theories and findings, but at the same time, it has to be acknowledged by the previous scholars by citing correctly. Without the rank of the wide galaxy of scholarly sources, ChatGPT may not have what it takes to produce a literature review that matches academic standards. Similarly, this lack of in-text citations renders the text unreliable or unscholarly because the facts and the ideas are no longer attributed to their sources. Then, ChatGPT technology may have content generation ability, but because it cannot access or incorporate previous literature, the need for human participation and intervention in helping academic writing processes is motivated.

### ***Methodology***

The continuous appearance of a mixed methods methodology, including surveys, interviews, and polls, in research topics for qualitative or quantitative studies discloses the drawback of the incapacity of ChatGPT to design research methods that fit specific research objectives. ChatGPT is outstanding in developing theoretical frameworks and models for different research contexts. However, its habit of proposing a mixed method presumes a one-size-fits-all product, which rules out the need for flexibility and adaptability in methodology development. The crucial factor here is that the research design and objectives should be the basis for choosing the appropriate data collection and research analysis methodologies to help address the research questions or hypotheses. Hence, though ChatGPT plays an important part in presenting the theoretical background information and the models related to it, human intervention and experience are paramount to helping researchers use the right methodologies that correspond with each study's unique purposes and objectives.

### ***Data Analysis***

The agility of ChatGPT in qualitatively and quantitatively inspecting data, whether the word count is short or long, is undoubtedly worth mentioning. Although ChatGPT can be a source of first-hand analysis and insight, it is still a machine and a human researcher will always need to check, verify, and amplify its work. The analysis provided by ChatGPT could be a very strong foundation. However, since the tool has word count limitations and can simplify details in a way, additional analysis and dimensionality are essential to make the results complex and precise. Human researchers can take advantage of the functions of ChatGPT to conduct an early analysis and get a preview of the information. However, a manual analysis is necessary to provide confidence in the outcomes' accuracy, depth, and completeness. So, although ChatGPT finds it effective to investigate data within a limited number of words, humans still play a great role in overseeing and collaborating for complete and accurate data analytics for the research and theory-based analyses.

### ***Findings***

The aptitude of ChatGPT to produce findings on keywords extracted from a given prompt is demonstrated. However, it may face the problem of tailoring the findings for a specified type of analysis. In that sense, it can only produce conclusions based on the information given; however, it cannot venture into detailed data analyses and produce very specific findings. The inability to fully comprehend the nuances of the information highlights the place of a human being in the interpretation and expertise in extracting information from multiple datasets. Human researchers should use ChatGPT designed preliminary

conclusions as a starting point but they must be cautious in further analysing and interpreting the data for specific and actual findings. So even though ChatGPT can provide initial support for the investigation, humans' role in narrowing and putting them into the context of the whole research must be addressed.

### **Conclusion**

ChatGPT is well-equipped to develop conclusions from prompted inputs. However, these outcomes could have generalist results when they are needed appropriately for specific research goals while promoting the effectiveness of the study. With a clear research objective, ChatGPT may gain direction in shaping its conclusions to the extent consistent with specific data analysis and findings. Accordingly, the proposed theory may be more general. It is a broader view of the topic and not a particular theory based on the research results. To achieve closure that reflects the research goals and findings, researchers must provide clear information and context to ChatGPT. With the incorporation of specific research objectives into the prompts, researchers can promote the creation of more targeted and relevant conclusions that pack in the key points arising from the data analysis process. ChatGPT can provide appropriate assistance at the final stages. However, human involvement and judgment will still be indispensable for aligning the findings with the specific objectives and outcomes of the study.

### **Recommendations**

Generating recommendations for research studies by ChatGPT is beneficial as it provides a general outlook. However, the recommendations could be more extensive as the model is constrained by information in its data. Knowledge of ChatGPT in September 2021 may cause it to miss the latest research outputs, resulting in the generation of advice that might be considered outdated. Moreover, the documents can only provide some details or need to be more specific or generalized to the current research field. Since new research is discovered every day, it is not advisable to rely solely on ChatGPT, which may not provide the newest and most exhaustive guidelines required by researchers. Hence, although ChatGPT may initially provide general suggestions, researchers should enrich this with curated literature reviews and specialists' opinions so that its recommendations remain consistent with the latest research findings and breakthroughs. Humanity, with expertise, skill, and a desire to discover new findings, continues to drive successful research recommendations in research studies.

### **CRAAP Test Analysis**

**Currency:** Content crafted by ChatGPT is session-generated from the prompts it responds to. This eliminates the currency problem because it does not cover any specific events or developments.

**Relevance/Reliability:** ChatGPT speaks in a voice that is influenced by the input it receives from users. Although it attempts to inform, it may not be able to distinguish between the relevancy and reliability of the sources it uses. Also, the fact that ChatGPT responses are from its training and algorithms reduces the accuracy of the generated content.

**Authority:** ChatGPT does not have an authoritative nature like other humans since the data it has been fed forms the basis of its activities. The capacity of AI-created content is directly dependent on the quality and credibility of the data incorporated by both the training and learning phases. Being unable to see the exact sources prevents one from perceiving the validity of the information provided.

**Accuracy:** ChatGPT is fed and generates its responses by patterns in its training data; hence, the correctness of this content depends on that data's quality. Although it can produce coherent and grammatically correct responses, it may be difficult to identify errors or inconsistencies in the information provided. When the provided information contains such problems, the training data might need to be fixed or biased.

**Purpose:** The role of ChatGPT is of a chatbot that responds to the input it receives, and its main function is to provide the user with relevant and succinct information. Nevertheless, it incorporates only unplanned or non-comprehending action, which differs from that of conscious decision-making because due to the

fact that its responses are based on patterns in the data upon which it is trained rather than a rational understanding.

### **Overall Analysis**

The overall analysis of the insight provided by the ChatGPT model gives a picture of a mixture of the capacity and the limitation of the model across the different research tasks of academia. ChatGPT proves itself to be an expert in the process of proposing research topics, writing introductions, analysing data, and conclusion- and recommendation-building. While there are certain drawbacks to the work performed by AI that require human hands, there is still a lot to be done before AI can replace human intervention and oversight.

Initially, ChatGPT does a great job of coming up with research topics, but it fails to summarize research and provide vital elements such as methods of research, data analysis, and findings. However, there is no citation in the introductions and literature reviews; such a move violates the academic integrity of the content, indicating a weakness in the ChatGPT system's ability to abide by academic guidelines.

Moreover, the ability of ChatGPT to frame general research objectives and default to the mixed methods approach in researching techniques, regardless of the study's qualitative or quantitative, shows the need for human refinement and precision. Also, the fact that it cannot interpret the analysis results as they are relevant to specific datasets reiterates the critical role played by human analysis and understanding to discern substantial conclusions from research data.

However, it has some limitations in processing data within a limited word count and carrying out precise analysis and recommendations according to written queries. Such capabilities can still be useful for researchers who can generate written outputs later. Nevertheless, it is important to consider the need for human editing, polishing, and inclusion of contextual information to validate the legitimacy, relevance, and academic standard of the regenerated content. According to the CRAAP test, crafted content shows how dependable it can be, depending on the level of relevance, reliability, accuracy, authority, and purpose. Even though it can do an exemplary job at passing on the information that the system receives, the limitations in the ability to differentiate between information that is factually true or not and the absence of any sense of intentionality of the results highlight potential pitfalls that can be associated with the reliability and accuracy of the produced content. Besides, because of the need for more transparency of the sources consulted in its training, discrimination of the authority of the data presents itself as a challenge. ChatGPT lays a solid foundation of practicality as a source for creating academic content and research. However, various limitations and the role of human expertise and evaluation are equally important.

### **Findings**

Content analysis crafted by ChatGPT showcases its capabilities and awareness of academic research problems but also has several weaknesses. ChatGPT proves itself relevant by proposing strong research ideas and processing initial data, outcomes, and recommendations. Nevertheless, several shortcomings demonstrate themselves in it; for example, the absence of essential components in abstracts, lack of references in the introduction and literature review, and only general formulations of research objectives. Also, ChatGPT specifiers that employ a mixed methods approach and the limitation in making the specific finding that will be used in the abstract data knowledge are indicators of the human role in the research process. However, irrespective of how helpful the ChatGPT is, human checks, additional refinement, and dissemination of information are required to ensure that the generated content is precise, pertinent, and academically accurate.

This CRAAP Test evaluation also captures the pros and cons of using ChatGPT as a source of information. The quality of the responses generated by the ChatGPT is current and relevant as it keeps updating the knowledge and information it contains in response to the users' input. However, questions of authenticity, expertise, correctness, and intent are attached to the content it generates. As the ChatGPT is based on its



training data and algorithms, the trustworthiness and accuracy of their output responses rely on the sources used in the training, which can result in errors and biases that can impact the reliability of the data. Moreover, human beings cannot verify the source of information the AI chatbot uses, which may put the accuracy of the information at risk. Hence, the users are cautioned to use ChatGPT as a supplement rather than as the main source of information while providing reasonable and coherent responses.

## **Discussion**

ChatGPT Generated Content Analysis's findings indicate the tool's ability to produce research-oriented content, e.g., the tool offers suggestions on research topics, creates introductions, and explores data. However, there were also some disadvantages, such as the failure to provide enough details about the research objectives, the literature reviews, and the recommendations that need to be specified more. These barriers, therefore, emphasize the need for humans to better the content generated by the ChatGPT. Although the machine can do some complicated tasks, like data processing within a limited word count, it cannot go beyond the basic algorithms it is programmed with to provide nuanced and contextually relevant outputs. Thus, although ChatGPT can be proven as a handy tool for quick research concepts, researchers still have to get further than merely its output as they need to add their values to scholarship in terms of validity, accuracy, and significance.

CRAAP Test analysis findings demonstrate the advantages and shortcomings of ChatGPT as a source of information. However, the tool highly scales up in providing pertinent, recent, and prompt answers to the user's request but calls for the reliability, correctness, authority, and purpose of the machine's content. It is not without question that the authenticity and accuracy of the output of ChatGPT are brought up, given that much of its training is based on the data and algorithms, the sources of which may, in some cases, contain biases and errors. The tool also needs to gain the skill to tell the reliability of the sources it addresses. Therefore, this tool might cause the loss of credibility of the information provided. This information tells us that the main users of ChatGPT should be careful about checking the output and not only rely on it but should use our other sources in case they doubt the correctness of the information they receive. ChatGPT is capable of providing this help by creating responses that are logical and relevant. However, human intervention is still necessary to verify and improve the information developed.

## **CONCLUSIONS AND RECOMMENDATIONS**

In the final analysis, the research aim was to check the efficacy of the ChatGPT language model in producing research work that is research-oriented within the English language and literature domain is a clear indicator of how much scholars can use AI-driven tools. ChatGPT shows competence in writing content concerning various forms of research that could be done, such as proposing a research topic and or collating data. However, these could be improved in their precision and citation accuracy. This allows human supervision to double-check the reliability and accuracy of the generated information before it can be used in real research work. Even with these constraints, ChatGPT will have its role as a tool to jump-start your research and some aspects of the writing process. Considering the purpose of the analysis to examine the quality, consistency, and relevance of materials generated by ChatGPT as articles for research, the results demonstrate both the strengths and the flaws of applying AI language models in a scholarly setting. The ChatGPT program can provide comprehensible and relevant answers to the prompts the users pose. However, its reliability and accuracy have been brought into question by some scholars, especially regarding the fact that it does not mention any irrelevant sources of information. Also, it is sometimes difficult for ChatGPT to understand the context of the user's question, which leads to confusion and debate. Therefore, the evidence reveals the need for a rigorous assessment of the quality of ChatGPT output by human scientists that will be used in research studies for the sake of the integrity and credibility of the research. In further research, collaboration between human experts and artificial intelligence-technological machines is important so that the AI language model can be fully exploited in academic work and its inherent flaws addressed.

Based on the findings of the study evaluating the efficacy and the quality of ChatGPT-generated research-oriented content, several recommendations emerge to enhance its relevance in academic research: Researchers who use ChatGPT should be vigilant and develop critical thinking skills to avoid the pitfalls of inaccuracy in the textual content. They should also ensure the content is checked rigorously through a manual review and validation process. Making quick modifications that trigger ChatGPT to be more specific and detailed in its answers is one way to address other possible shortcomings of the AI bot. Training data ChatGPT and algorithms will demonstrate constant improvement as time passes, leading to the greater ease of the system, generating scholarly content that coincides with the academic levels. The researchers can also use connectors manually to make their writings well connected; to render an air of flow of ideas. Moreover, researchers need to be aware and exploit the potential of ChatGPT not only as a tool that serves the inquisitive aspect of research but also to speed up some parts of the scholarly writing process; at the same time, they should not forget about the need for human supervision and interaction to promote the reliability and verification of the generated information.

### **Implications**

The implication of the study, which compared the capability of ChatGPT with that of humans while generating research-related details within the English language and literature domain, is vast. It points out that the rise of AI will redefine the horizons of academic thought, similarly, it might open more opportunities for scholars in future but it will have its fetters, too. ChatGPT, undoubtedly, illustrates a remarkable skill in suggesting research topics and producing frameworks; its incapability in abstract writing, citation incorporation, and methodology adaptation shift the platform toward humans to ensure that everything, including research outputs, is done with accuracy, relevance and precision. However, the study points out that researchers should approach AI tools like ChatGPT with a critical mindset. They need to understand power of the tool in the research process and, at the same time, recognize that AI also has limitations; what the tool can do and what cannot, meaning by there still remains a need for human input and validation of the generated data. Ultimately, the study highlights the need for reflection on the evolving dynamics between AI technology and human scholarship, emphasizing the significance of collaboration and discernment in leveraging AI capabilities for academic inquiry and advancement.

### **REFERENCES**

- Amirjalili, F., Neysani, M., & Nikbakht, A. (2024). Exploring the boundaries of authorship: a comparative analysis of AI-generated text and human academic writing in English literature. In *Frontiers in Education* (Vol. 9, p. 1347421). Frontiers Media SA.  
<https://doi.org/10.3389/feduc.2024.1347421>.
- Blakeslee, S. (2004). The CRAAP test. *Loex Quarterly*, 31(3), 4.
- Chaka, C. (2023). Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: The case of five AI content detection tools. *Journal of Applied Learning and Teaching*, 6(2), 1-11.  
<https://doi.org/10.37074/jalt.2023.6.2.12>
- Esparrago-Kalidas, A. J. (2021). The effectiveness of CRAAP test in evaluating credibility of sources. *International Journal of TESOL & Education*, 1(2), 1-14.
- Essberger, J. (2021). The Winepress. EnglishClub.com. <https://www.englishclub.com/reading/story-winepress.php>.
- Haleem, A., Javaid, M., & Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *Bench Council transactions on benchmarks, standards and evaluations*, 2(4), 100089. <https://doi.org/10.1016/j.tbench.2023.100089>
- Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific Reports*, 13(1), 18617.  
<https://doi.org/10.1038/s41598-023-45644-9>.

- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), ep464. <https://doi.org/10.30935/cedtech/13605>.
- Khan, N. M. (2023). Implementation of Inquiry-Based Learning in Pakistani Government Colleges: An Exploratory Study. *Journal of Asian Development Studies*, 12(4), 796-807.
- Khlaif, Z. N., Mousa, A., Hattab, M. K., Itmazi, J., Hassan, A. A., Sanmugam, M., & Ayyoub, A. (2023). The potential and concerns of using AI in scientific research: ChatGPT performance evaluation. *JMIR Medical Education*, 9, e47049. <https://doi.org/10.2196/47049>.
- Kolade, O., Owoseni, A., & Egbetokun, A. (2024). Is AI changing learning and assessment as we know it? Evidence from a ChatGPT experiment and a conceptual framework. *Heliyon*. <https://doi.org/10.1016/j.heliyon.2024.e25953>.
- Liang, W., Izzo, Z., Zhang, Y., Lepp, H., Cao, H., Zhao, X., & Zou, J. Y. (2024). Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews. *arXiv preprint arXiv:2403.07183*. <https://doi.org/10.48550/arXiv.2403.07183>.
- Mindner, L., Schlippe, T., & Schaaff, K. (2023). Classification of human-and ai-generated texts: Investigating features for chatgpt. In *International Conference on Artificial Intelligence in Education Technology* (pp. 152-170). Singapore: Springer Nature Singapore. [https://doi.org/10.1007/978-981-99-7947-9\\_12](https://doi.org/10.1007/978-981-99-7947-9_12).
- Nazeer, I., Khan, N. M., Nawaz, A., & Rehman, J. (2024). An Experimental Analysis of Pragmatic Competence in Human-ChatGPT Conversations. *Pakistan Journal of Humanities and Social Sciences*, 12(1), 424-435. <https://doi.org/10.52131/pjhss.2024.v12i1.2061>.
- Nazeer, I. (2022). *Love Gestures*. LULU Publishers USA. <https://www.lulu.com/shop/imran-nazeer/love-gestures/paperback/product-r9rgnn.html?page=1&pageSize=4>.
- Niloy, A. C., Akter, S., Sultana, N., Sultana, J., & Rahman, S. I. U. (2024). Is Chatgpt a menace for creative writing ability? An experiment. *Journal of computer assisted learning*, 40(2), 919-930. <https://doi.org/10.1111/jcal.12929>.
- Rafique, H., Nazeer, I., & Rehman, J. (2024). The Impact of ChatGPT on Language Evolution: A Linguistic Analysis. *Journal of Education and Social Studies*, 5(1), 56-68. <https://doi.org/10.52223/jess.2024.5106>.
- Reed, P. A. (2023). *Is ChatGPT Creative? Cognitive-affective Responses to AI-generated Stories* (Doctoral dissertation, Fielding Graduate University).
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>.
- Tirado-Olivares, S., Navío-Inglés, M., O'Connor-Jiménez, P., & Cózar-Gutiérrez, R. (2023). From Human to Machine: Investigating the Effectiveness of the Conversational AI ChatGPT in Historical Thinking. *Education Sciences*, 13(8), 803. <https://doi.org/10.3390/educsci13080803>.
- Wang, Y., Pan, Y., Yan, M., Su, Z., & Luan, T. H. (2023). A survey on ChatGPT: AI-generated contents, challenges, and solutions. *IEEE Open Journal of the Computer Society*. <https://doi.org/10.1109/OJCS.2023.3300321>.
- Young, J. C., & Shishido, M. (2023). Investigating OpenAI's ChatGPT Potentials in Generating Chatbot's Dialogue for English as a Foreign Language Learning. *International Journal of Advanced Computer Science and Applications*, 14(6), 65-72. <https://doi.org/10.14569/IJACSA.2023.0140607>.