

# Meida Rabia - Examining the Validity

*by - -*

---

**Submission date:** 15-Dec-2023 12:19PM (UTC+0200)

**Submission ID:** 2189978866

**File name:** Meida\_Rabia\_-\_Examining\_the\_Validity.docx (60.61K)

**Word count:** 3800

**Character count:** 23071

# Examining the Validity and Reliability of ChatGPT 3.5-Generated Reading Comprehension Questions for Academic Texts

DOI: <https://doi.org/10.47175/rielsj.v1ix.xx>

| Meida Rabia Sihite<sup>1,\*</sup> | Meisuri<sup>2</sup> |

<sup>1</sup>English Education Study Program, Universitas Alwashliyah, Medan, Indonesia

<sup>2</sup>Postgraduate Program, Universitas Negeri Medan, Medan, Indonesia

[meidarabia55@gmail.com](mailto:meidarabia55@gmail.com)  
[meisuri@yahoo.com](mailto:meisuri@yahoo.com)

## ABSTRACT

This research examines the capacity of ChatGPT 3.5 in generating reading comprehension questions for academic texts, with a focus on their alignment with higher-order cognitive skills as per Bloom's Taxonomy. A paper-based test comprising 30 multiple-choice questions was constructed using ChatGPT 3.5, based on three selected TOEFL ITP reading comprehension passages. The study employed a mixed-methods approach, integrating qualitative content analysis to assess the cognitive level of each question and quantitative methods to analyze student responses. Data collection involved administering the AI-generated questions to students and scoring their responses. Analysis techniques included Pearson correlation coefficients to determine validity and reliability analysis using Cronbach's Alpha to measure internal consistency. The findings revealed that ChatGPT 3.5 is capable of producing questions that cover a range of cognitive levels, from analysis to creation, however only 10 out of 30 questions met the validity criteria, indicating a need for improvement in the AI's question generation process. The reliability of these questions was moderate, suggesting a reasonable level of internal consistency. The study concludes that while AI-generated questions show promise in educational assessments, ongoing improvement of AI models is necessary to enhance their effectiveness. The implications of this research are significant for the future integration of AI in educational settings, indicating a potential role for AI in developing meaningful assessment tools. The study recommends future research to explore various question types and incorporate student feedback to optimize the effectiveness of AI in education.

## KEYWORDS

ChatGPT 3.5; validity; reliability; reading comprehension questions

## INTRODUCTION

The integration of AI language models, such as ChatGPT 3.5, in educational and language education domains has garnered significant attention due to their potential to revolutionize the creation of educational materials and language assessments. These advanced models have demonstrated utility in generating virtual patient simulations, quizzes, and educational materials, showcasing their potential to enhance learning experiences (Eysenbach, 2023; Hung et al., 2023). However, concerns have been raised regarding the reliability and validity of AI-generated content, especially when applied to assess reading comprehension proficiency in academic texts (Rahman & Watanobe, 2023; Tyson, 2023; Vandiver, 2008).

The existing literature has highlighted the potential of AI in education, with studies demonstrating the increasing utility of AI as a national growth engine and its potential to provide tremendous value across various educational fields (Su & Yang, 2023). The use of artificial intelligence (AI) in education has gained significant attention in recent years, with a growing number of educational institutions and organizations exploring the potential

# Meida Rabia - Examining the Validity

## ORIGINALITY REPORT

15%

SIMILARITY INDEX

13%

INTERNET SOURCES

12%

PUBLICATIONS

%

STUDENT PAPERS

## PRIMARY SOURCES

1	<a href="http://files.eric.ed.gov">files.eric.ed.gov</a> Internet Source	4%
2	<a href="http://www.researchgate.net">www.researchgate.net</a> Internet Source	1%
3	<a href="http://eprints.binadarma.ac.id">eprints.binadarma.ac.id</a> Internet Source	1%
4	<a href="http://pdffox.com">pdffox.com</a> Internet Source	1%
5	<a href="http://core.ac.uk">core.ac.uk</a> Internet Source	1%
6	<a href="http://moam.info">moam.info</a> Internet Source	1%
7	<a href="http://mededu.jmir.org">mededu.jmir.org</a> Internet Source	1%
8	<a href="http://vtechworks.lib.vt.edu">vtechworks.lib.vt.edu</a> Internet Source	1%
9	<a href="http://acikbilim.yok.gov.tr">acikbilim.yok.gov.tr</a> Internet Source	1%

10	<a href="http://journal.unhas.ac.id">journal.unhas.ac.id</a> Internet Source	<1 %
11	<a href="http://www.oer.unn.edu.ng">www.oer.unn.edu.ng</a> Internet Source	<1 %
12	Zied Bahroun, Chiraz Anane, Vian Ahmed, Andrew Zacca. "Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis", <i>Sustainability</i> , 2023 Publication	<1 %
13	Jussi S. Jauhiainen, Agustín Garagorry Guerra. "Generative AI and ChatGPT in School Children's Education: Evidence from a School Lesson", <i>Sustainability</i> , 2023 Publication	<1 %
14	"Innovative Technologies and Learning", Springer Science and Business Media LLC, 2023 Publication	<1 %
15	<a href="http://brieflands.com">brieflands.com</a> Internet Source	<1 %
16	<a href="http://www.593dp.com">www.593dp.com</a> Internet Source	<1 %
17	<a href="http://www.mdpi.com">www.mdpi.com</a> Internet Source	<1 %

18

Krithi Pushpanathan, Zhi Wei Lim, Samantha Min Er Yew, David Ziyou Chen et al. "Popular large language model chatbots' accuracy, comprehensiveness, and self-awareness in answering ocular symptom queries", iScience, 2023

Publication

<1 %

19

Vijay Kanabar. "Chapter 30 An Empirical Study of Student Perceptions When Using ChatGPT in Academic Assignments", Springer Science and Business Media LLC, 2023

Publication

<1 %

20

Xinghua Wang, Linlin Li, Seng Chee Tan, Lu Yang, Jun Lei. "Preparing for AI-enhanced education: Conceptualizing and empirically examining teachers' AI readiness", Computers in Human Behavior, 2023

Publication

<1 %

21

[link.springer.com](https://link.springer.com)

Internet Source

<1 %

22

[www.iejee.com](https://www.iejee.com)

Internet Source

<1 %

23

Supawadee Suppadungsuk, Charat Thongprayoon, Pajaree Krisanapan, Supawit Tangpanithandee et al. "Examining the Validity of ChatGPT in Identifying Relevant Nephrology Literature: Findings and

<1 %

# Implications", Journal of Clinical Medicine, 2023

Publication

---

24

[hdl.handle.net](https://hdl.handle.net)

Internet Source

<1 %

---

---

Exclude quotes      Off

Exclude matches      Off

Exclude bibliography      On