

## Comment: Importance of critical thinking to understand ChatGPT

Frank Emmert-Streib, Predictive Society and Data Analytics Lab, Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland  
Email: v@bio-complexity.com

This is a response to [1] providing constructive criticism about [2,3]. In the following, we will address the points of criticism in the order of appearance as they were presented in [1].

A point of criticism raised by [1] seems to be about the term “understanding”. The authors note that ChatGPT does not have the capacity for self-awareness or cognition nor is ChatGPT conscious in any form. Furthermore, the authors point out that ChatGPT is based on predictions and not understanding. In our opinion this is correct and none argued otherwise. In fact, the term “understand” occurs only once in [3], namely in the title, stated as a question - with a question mark at the end. Also the study in [2] does not aim to go beyond the predictive ability of ChatGPT.

Furthermore, the authors write [1]: “The authors interpret the language model's performance as indicative of an "understanding" of genetics.” This statement is not correct, instead we wrote [3]: “This creative usage allowed to show that in comparison to a heterogeneous group of humans with an uncontrolled background, ChatGPT performed similarly with an overall accuracy of almost 70%.” The important words here are “performed” and “accuracy”, both refer to a statistical evaluation of the analysis rather than any cognition-related term mentioned above.

For the interested reader, we would like to point out that if anyone would intend to go along this road, one would need to talk about a *Turing test* (or the imitation game) [5] and *artificial general intelligence* (AGI) [6]. Neither has been the topic in [3].

The next is a reiteration of a similar sentiment [1]: “Hence, it is an inaccurate stretch to attribute a score of 70% correct answers to a genuine understanding of genetics.” Again, this was not an argument that can be found in [3]. From a statistical perspective, we would also like to add that it is well known that one data set alone does not allow to estimate the *expected generalization error* reliably of a prediction model, as ChatGPT [4]. This point relates to the theoretical foundations of statistics whereas the expected generalization error is a population estimator of the ideal (theoretical) predictive performance of a model. The connection to [2] is that there only one dataset for a particular task has been studied.

The next set of critics expresses concerns regarding the utilization of ChatGPT in education. However, upon reviewing the statements presented in [1], it becomes evident that no clearly articulated point specifically addressing this concern can be found. For this reason, no particular response can be provided. However, we would like to highlight that nobody has argued that ChatGPT, in its current form, will replace teachers or professors anytime soon. Nonetheless, even in its present state, ChatGPT can be a valuable tool for learning.

It is clear that effective learning requires active engagement in critical and analytical thinking, rather than simply adhering to the guidance of either a human teacher or ChatGPT. In reality, no perfect teacher exists, which means that learning inherently occurs in the presence of an imperfect teacher. Furthermore, the development of critical thinking skills requires exposure

to incorrect statements or ideas to practice the evaluation of statements [7]. Without such exposure, the ability to question and critically evaluate sources of information remains limited. By encountering and examining incorrect statements or ideas, individuals can enhance their capacity to discern reliable information and bolster their critical thinking abilities. It is crucial to avoid adopting a black-and-white perspective in this regard and instead strive for a balanced approach that considers all relevant sources of learning. It is worth acknowledging that ChatGPT, while an imperfect source of information, is not random in its responses. As a result, it provides users with opportunities to demonstrate critical thinking

We highly encourage those who have not yet done so to explore the application of ChatGPT in fields such as genetics or any other domain of interest. While not every answer provided will be correct or even useful some will be. It is this selective usefulness of ChatGPT that we find beneficial and worth exploring. This is expressed in the following statement [3]: “Certainly, even in its current form ChatGPT can be used for this purpose but the learner requires a high level of maturity by not taking every answer as ground-truth but as a suggestion that needs to be checked.” Since [2] is published in the European Journal of Human Genetics which is a research journal about human genetics and genomics the educational level of the readers is advanced allowing to assume that the readers are familiar with critical thinking.

Being inspired by the raised arguments, we believe that the issue at hand is not so much a point of disagreement but rather the result of misunderstanding. While ChatGPT can be used by anyone and in any field, its understanding requires expertise in a number of technical fields. Specifically, for the algorithms behind ChatGPT proficiency in natural language processing (NLP) and artificial intelligence (AI) is essential for understanding and implementing its intricate workings. Conversely, to assess its performance, a solid foundation in statistics is necessary to analyze and interpret the results accurately. Furthermore, when interpreting the answers to questions, domain-specific knowledge about the field relevant to the questions becomes indispensable. In other words, this is a typical clash in multi-disciplinary research where each discipline tries to understand the problem exclusively within its own domain. However, this does not work.

ChatGPT is a good example to emphasize the importance of interdisciplinary education. In this context, data science provides a good foundation to learn about all required skills extending over artificial intelligence, machine learning and statistics [8].

Finally, we would like to note that we do not see where ChatGPT has been hyped. Instead, the spirit we see it may be best described by a quote often attributed to George Box [9]: “All models are wrong, but some are useful.” This quote encapsulates the foundational principles applicable to all scientific endeavors, providing a solid grounding for the pursuit of knowledge and discovery.

We would like to end this brief response by thanking the authors for being critical and we would also like to encourage the readers to raise questions and to challenge statements not only about ChatGPT and AI but in general because a critical discourse is what defines science not consensus [10].

### **Competing interests**

The authors declare no competing interests.

## Author Contributions

FES conceived and drafted the manuscript.

## Funding

No funding.

## References

1. Partha Pratim Ray and Poulami Majumder. Beyond the Hype: A Comprehensive Critique of ChatGPT's Applicability in Genetics, *European Journal of Human Genetics*, 2023.
2. Dat Duong and Benjamin D Solomon. Analysis of large-language model versus human performance for genetics questions. *European Journal of Human Genetics*, 2023.
3. Frank Emmert-Streib. Can ChatGPT understand genetics? *European Journal of Human Genetics*, 2023.
4. Frank Emmert-Streib, and Matthias Dehmer. "Evaluation of regression models: Model assessment, model selection and generalization error." *Machine learning and knowledge extraction* 1, no. 1 (2019): 521-551.
5. Alan M. Turing, I.—COMPUTING MACHINERY AND INTELLIGENCE, *Mind*, Volume LIX, Issue 236, October 1950, Pages 433–460, <https://doi.org/10.1093/mind/LIX.236.433>
6. Ben Goertzel. "Artificial general intelligence: concept, state of the art, and future prospects." *Journal of Artificial General Intelligence* 5, no. 1 (2014): 1.
7. Facione, P.A., Sanchez, C.A., Facione, N.C. and Gainen, J., 1995. The disposition toward critical thinking. *The Journal of General Education*, 44(1), pp.1-25.
8. Frank Emmert-Streib, Salissou Moutari and Matthias Dehmer, 2022. *Mathematical foundations of data science using R*. Walter de Gruyter GmbH & Co KG.
9. George E. Box. 1976. Science and statistics. *Journal of the American Statistical Association*, 71(356), pp.791-799.
10. Polycarp Ikuenobe. 2001. Questioning as an epistemic process of critical thinking. *Educational Philosophy and Theory*, 33(3-4), pp.325-341.