

SOLVING A PROBLEM WITH AI: CONSIDERATIONS ON TEACHING AND LEARNING MATHEMATICS

Camilla Spagnolo¹, Giorgio Bolondi¹, Alessandro Gambini²

¹*Free University of Bozen/Bolzano, Camilla.Spagnolo@unibz.it and Giorgio.Bolondi@unibz.it*

²*La Sapienza, University of Rome, alessandro.gambini@uniroma1.it*

The application of artificial intelligence (AI) in STEM education and particularly in mathematics, as an emerging field, is confronted with a challenge of integrating diverse AI techniques and complex educational elements to meet instructional and learning needs. We discuss a possible use of ChatGPT for investigating teachers' practices and beliefs.

Keywords: artificial intelligence, chatGPT, perception, mathematics education, computer science.

INTRODUCTION

Since January, chatGPT has entered massively into the communication of major newspapers (Aydın & Karaarslan, 2022), even though the chatbot in question (in the first beta version) was released by OpenAI manufacturing company on November 30, 2022. After remarks by Microsoft CEO Nadella in Davos, it is known that chatGPT will be integrated into many Microsoft products primarily Bing the company's flagship browser. The impact undoubtedly of the new tool has been strong, the final product is awaited.

The impact will involve all sectors of our society, but for now the most direct seems to be the training and cheating effect.

ChatGPT, at the moment, is not certified and does not cite sources where it gets its information. A central and strategic question arises: what will be the impact on educational systems, specifically in the teaching of science disciplines and mathematics in particular? Can teaching not take into account the new perspectives and challenges that AI poses, is one prepared for this "challenge"?

The EU recently drafted an interesting document "Ethical guidelines for educators on the use of artificial intelligence (AI) and data in teaching and learning" (European Commission, 2022).

Artificial intelligence in education (AIEd) is an emerging interdisciplinary field that applies AI technologies in education to transform and promote the instructional and learning design, process and assessment (Chen et al., 2020; Holmes et al., 2019; Hwang et al., 2020; Xu & Ouyang, 2022).

There are currently few in-depth studies on the topic of artificial intelligence in mathematics education (Knill et al., 2004; Garrido, 2012).

Artificial intelligent projects like "Eliza" (Weizenbaum, 1966), "Student" or "GPS" (Russel, 2010) attempted to produce agents which can solve general mathematical problems. One soon realized that this goal was too ambitious. Not only because of technical hurdles, but also because one had to realize that many problems are computationally hard and can not be settled in reasonable time. While this slowed the momentum of AI developments in education, it did not stop it.

ITALIAN CONTEXT AND METHODOLOGY OF THE STUDY

In Italy there are two cycles of education (www.miur.gov.it). The first cycle of education consists of two consecutive and mandatory school courses: primary school (lasting five years, for students aged 6 to 11 that corresponds to grade 1 to 5), and middle school (lasting three years, for students aged 11 to 14 that corresponds to grade 6 to 8). After finishing middle school, students have access to the second cycle of education, which ends at age 19 (grade 9 to 13). Mandatory education lasts 10 years (from age 6 to 16), and includes the eight years of the first cycle of education and the first two years of the second cycle.

In addition, in the Italian context, we have the possibility to track some students' difficulties over time thanks to INVALSI tests (tests with the purpose of measuring students' levels of competence in relation to the Italian curricular Guidelines) which were administered since in 2008 in grades 2, 5, 8, 10 and 13 from the National Institute for the Evaluation of the Educational System (from 2009 to 2013, the tests also covered grade 6).

Up to now, the Italian Ministry of Public Education has established the standardised assessment of the Italian educational system, and commissioned the INVALSI (www.invalsi.it) to carry out annual surveys nationwide to all students in the second and fifth classes of primary school (grades 2 and 5), middle school third class (grade 8), and high school (grades 10 and 13). The INVALSI Institute carries out periodic and systematic checks on students' knowledge and skills (about reading comprehension, grammatical knowledge and mathematical competency), and on the overall quality of the educational offers from schools and vocational training institutions; in particular, it runs the National Evaluation System (SNV). The INVALSI standardised tests were created for system evaluation and this is their primary purpose. The tests are administered every year at census level and student results are provided to each school institution. Results and questions of the INVALSI tests are considered as a resource also for researchers in the field of mathematics education (Garuti & Martignone, 2015) and are used in national and international research (e.g. Spagnolo et al., 2021).

The SNV Framework is designed taking into consideration the Italian National Guidelines, in which argumentation is considered between the competence goals for every school grade (Garuti & Martignone, 2019). For this reason it is possible to find mathematical tasks that require the recognition or production of a justification, as the tasks we selected for our questionnaire design.

In particular, the tasks chosen in this study are INVALSI tasks; this choice has been made because it allows us to have some extra information that provides a further background for our study, such as the performance of Italian students (which is related to task peculiarities). In our study, we focused on grade 8 INVALSI task, which represent the last year of middle school in Italy.

We identified an INVALSI task (Figure 1) for grade 8 and selected some of the typical answers that grade 5 students provide.

D15. n is a natural number. Consider the statement, "If n is even then $n+1$ is a prime number." Is the statement true or false?

Choose the answer and complete the sentence.

The statement is true because

.....

.....

The statement is false because

.....

.....

Figure 1. Task D15, Math INVALSI Test Grade 8, 2016, www.gestinv.it

The same task was proposed to chatGPT and the response was as follows.

ChatGPT: The statement is false because, for example, when $n=2$, we have that n is even, but $n+1=3$ which is a prime number. However, if we take $n=4$, then n is even, but $n+1=5$ which is a prime number.

So, it is not possible to say that "if n is even then $n+1$ is a prime number" in general.

We asked 15 middle school teachers to assess all selected answers (including that of chatGPT). The assessment was done according to the following criteria: accuracy of content, accuracy of language, coherence to the task, coherence to the claim, completeness in calculation, completeness in argumentation.

Then it was revealed to the teachers that one of the solutions was made by chatGPT, asking them to recognize it.

We investigate whether this could be useful to bring out beliefs and practices of teachers.

DISCUSSION AND PRELIMINARY RESULTS

The data of the national Italian sample are gathered, analyzed and released by INVALSI. Data can be found at www.gestinv.it. Following, we report the national results of the task shown in Figure 1.

This task aims to assess the ability to use of some typical forms of mathematical thinking (conjecture, argue, verify, generalize); in particular supporting their convictions by bringing appropriate examples and counterexamples and using argumentation to justify their choices.

The percentage of correct answers of the Italian national sample is 37,1%, while the percentage of incorrect answers is 41.4% and of unanswered questions is 21.6%.

Below we show 4 categories of student answers (Figure 2).



Figure 1. Examples of student answers

The transcription of the 4 responses (Figure 2) to the task in Figure 1 follows.

Student a: The statement is false because if we add 1 it is no longer a prime number.

Student b: The statement is false because not all prime numbers are odd.

Student c: The statement is true because if to an even number you add 1 it is normal that the result will be a number 1. [Answer deleted]

The statement is false because it is not true that the result will necessarily be an even number.

Student d: The statement is true because if to an even number you add one it is normal that the result will be a prime number. [Answer deleted]

The statement is false because it is not true that the result will necessarily be a prime number.

At the same task chatGPT answers by giving examples (not counterexamples) and without maintaining consistency with the request of the task and its own claim.

The study is a work in progress, but all 5 protocols will be assessed by 15 middle school teachers.

After assessing them, we will ask her to identify the chatGPT one. We believe this will bring out some teachers' beliefs, such as the idea that AI always gives a correct answer.

CONCLUSION

It is clear from the analysis that good and specific questions must be asked in order to obtain more and more significant answers to gain knowledge about the object of the research. If the question is generic, the answer is generic, often so much so that it is incorrect.

This suggests that the fears surrounding the use of chaptGPT by students, and which have led some U.S. colleges and universities to ban its use, are overblown: if the output of chaptGPT is complex, articulate, and robust, there is accurate, detailed research behind it, grounded in good, surgical questions.

This is one of the fundamental goals of teaching action, and that is, to help students ask good questions in order to actually access the knowledge that is only theoretically available to everyone now.

So two sticking points for students: learning to ask good questions and checking the reliability of the information obtained. It seems to us that this is enough to ensure the need for the presence of competent teachers capable of helping students grow as knowledgeable, critical and ethical subjects. AI can become a formidable ally of the teacher in his or her many activities, the figure of the teacher assumes, in this new context, a strategic role even at a social level. This is a necessary and different approach to knowledge.

Finally, we believe that this study also allows to highlight some teachers' beliefs, such as the idea that AI always answers correctly.

REFERENCES

- Aydın, Ö., & Karaarslan, E. (2022). OpenAI ChatGPT generated literature review: Digital twin in healthcare. *Available at SSRN 4308687*.
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Commissione europea, Direzione generale dell'Istruzione, della gioventù, dello sport e della cultura, (2022). *Orientamenti etici per gli educatori sull'uso dell'intelligenza artificiale (IA) e dei dati nell'insegnamento e nell'apprendimento*, Ufficio delle pubblicazioni dell'Unione europea. <https://data.europa.eu/doi/10.2766/438>
- Garrido, A. (2012). AI and mathematical education. *Education Sciences*, 2(1), 22-32.
- Garuti, R., & Martignone, F. (2015). The SNV (INVALSI) experience. Teaching and learning mathematics: resources and obstacles. *Proceedings of CIEAEM*, pp. 95-98.
- Garuti, R., & Martignone, F. (2019). Assessment and argumentation: an analysis of mathematics standardized items. In U. T. Jankvist, M. Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Eleventh Congress of the European Society for Research in Mathematics Education*, Freudenthal Group & Freudenthal Institute, Utrecht University and ERME, pp. 4075-4082.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100001. <https://doi.org/10.1016/j.caeai.2020.100001>
- Knill, O., Carlsson, J., Chi, A., & Lezama, M. (2004). An artificial intelligence experiment in college math education. *Preprint available at http://www.math.harvard.edu/~knill/preprints/sofia.pdf*.
- Russell, S. J. (2010). *Artificial intelligence a modern approach*. Pearson Education, Inc.

- Spagnolo, C., Capone, R., & Gambini, A. (2021). Where do students focus their attention on solving mathematical tasks? An eye tracker explorative study. Inprasitha, M., Changsri, N., & Boonsena, N. (Eds). (2021). *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education* (Vol.4). Khon Kaen, Thailand: PME, pp. 89-96.
- Xu, W., Ouyang, F. (2022). The application of AI technologies in STEM education: a systematic review from 2011 to 2021. *IJ STEM Ed* **9**, 59 (2022). <https://doi.org/10.1186/s40594-022-00377-5>
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36-45.