\*\*Is ChatGPT a Math Whiz?\*\*

\*Mathematics Education in the Age of AI\*

In the digital era, the way we teach and learn has drastically changed. Artificial intelligence (AI) is taking up an increasingly significant place in our daily lives, offering new possibilities to convey and acquire knowledge. A remarkable example of this is ChatGPT, a powerful language model (LLM, Large Language Model) developed by OpenAI, which operates through artificial neural networks designed to understand and generate human-like text. But what role can and should ChatGPT play in mathematics education? Can this AI truly be considered a math whiz and assist teachers in making lessons more effective, interesting, or even more differentiated?

In this article, we aim to provide educators with insights into the capabilities and limitations of ChatGPT in mathematics instruction. We examine how this technology works and how it solves mathematical problems. Simultaneously, we take a critical look at the challenges that must be considered when dealing with AI in the classroom. Can ChatGPT be a valuable assistant in mathematics education, or should traditional teaching methods continue to be preferred?

The current versions of ChatGPT, based on the GPT-4 architecture, offer various models that differ in their capabilities and areas of application:

- \*\*GPT-4o\*\*: This version is particularly efficient and powerful in solving complex mathematical problems and scientific calculations. It offers fast and precise results and is ideal for demanding teaching situations.

- \*\*GPT-4\*\*: The main version provides comprehensive understanding and high accuracy across a wide range of mathematical concepts. It is suitable for detailed explanations and in-depth analyses in the classroom.

- \*\*GPT-4-mini\*\*: This more compact version is resource-efficient and intended for mobile use or environments with limited computing resources. It can solve basic to moderately difficult mathematical problems and is ideal for flexible and easily accessible solutions in the classroom.

- \*\*GPT-3.5\*\*: The predecessor version, GPT-3.5, also offers strong mathematical capabilities, albeit with slightly lower accuracy and less contextual understanding compared to GPT-4. Nevertheless, it is a solid choice for many teaching scenarios and offers a free alternative to the latest models (as of August 2024).

In addition to simple text inputs, GPT can also process file types, for example, analyze and summarize texts. The response is usually output in text form but can also be provided in files (e.g., Office formats). The paid versions offer the ability to process larger amounts of data at once and to create images. Moreover, GPT-4o can be personalized for individual needs by particularly considering certain pre-loaded data sets. Additionally, users gain access to other published GPTs, including those that solve mathematical problems with the help of Wolfram Alpha.

\*\*How Does ChatGPT Work?\*\*

Initially, the use of ChatGPT in the mathematical field painted a sobering picture. Even with simple applications, such as calculating the zeros of functions, the solutions were usually not correctly computed. Instead, they were either statistically derived from the coefficients or were freely invented (hallucinated).

Nowadays, solving mathematical problems with ChatGPT is based on a multi-stage processing of text and mathematical symbols. The process begins with \*\*tokenization\*\*, a crucial step where the input text is broken down into smaller units called tokens. These tokens can be words, punctuation marks, or mathematical symbols. They are processed through the neural network of the model, which has been trained with millions of texts and mathematical tasks. The model recognizes patterns and applies mathematical rules, then performs the mathematical operations and generates an answer step by step. It can solve simple tasks directly and explain more complex problems in detail. When applying more complex mathematical methods or formulas, GPT accesses mathematical libraries in Python, for example, to solve the tasks. This clarifies which steps and calculations have been carried out. As previously mentioned, specialized GPTs can access other resources and achieve different results. Literature already exists that demonstrates how ChatGPT can be used to solve even complex mathematical problems (Helfrich-Schkarbanenko, 2023). However, precise instruction and the choice of mathematical tools are required.

\*\*How Well Does GPT Perform When Solving Tasks Independently?\*\*

OpenAI, the developer, initially investigated this question by having GPT take the mathematics section of the American Scholastic Assessment Test (SAT), a part of the college entrance exams. GPT-4 scored in the 89th percentile, significantly better than average. For comparison, Harvard students are usually at least in the 90th percentile. However, these results must be put into perspective. On one hand, it is not clear to what extent these tests were already used to train GPT, and on the other hand, the SAT contains comparatively simple, mostly calculation-heavy tasks.

Other studies have arrived at more differentiated results: performance depends on the mathematical level, the field, and the type of task. GPT is capable of solving simple standard tasks well but has increasing difficulties at higher grade levels. Apart from the difficulty level, context and especially the wording play a significant role in application tasks.

\*\*First Example (Spreitzer et al., 2024)\*\*

We present the following task to ChatGPT and analyze its solution:

\*(Task not provided in the text)\*

![Figure 1: Solution by ChatGPT 3.5](#)

\*Figure 1: Solution by ChatGPT 3.5\*

The solution approach using the geometric series (and the result as well) is incorrect. It is suspected that the wording "... he jumps half the distance ..." triggered GPT to use the geometric series, as halving distances is often required in application tasks involving geometric series. However, this problem does not occur with the new versions of ChatGPT.

\*\*Second Example: Tank Problem According to Blum\*\*

\*(Task not provided in the text)\*

![Figure 2: Solution of the Tank Problem by ChatGPT 4](#)

\*Figure 2: Solution of the Tank Problem by ChatGPT 4\*

The solution is mathematically correct; missing assumptions were correctly identified. However, this is a purely mathematical solution; other factors such as ecological or temporal considerations do not enter into the argumentation, even upon request.

\*\*Third Example: Traffic Jam Problem According to Maaß\*\*

\*(Task not provided in the text)\*

![Figure 3: Solution by ChatGPT 4](#)

\*Figure 3: Solution by ChatGPT 4\*

The solution is not wrong per se. GPT, as in the previous task, also identified missing information and made plausible assumptions. However, the solution assumes that the highway has only one lane.

\*\*Can ChatGPT Teach Mathematics Well?\*\*

Heterogeneity is considered one of the main challenges in many school classes and, in theory, would be best addressed in individual tutoring situations. The idea that students receive tasks from ChatGPT, GPT corrects them, and based on the solutions, creates and supervises individual learning paths sounds utopian given the aforementioned deficits. However, this has not yet been conclusively researched, and indeed new, at least sequential, application fields seem to be opening up. In early secondary education, for example, rational and decimal numbers are introduced. Certain calculation errors, so-called systematic errors like \( \frac{3}{5} + \frac{4}{7} = \frac{7}{12} \), when repeated, suggest comprehension problems. From version 4 onwards, GPT is quite capable of identifying such error patterns in scanned student solutions and providing appropriate suggestions for remedial tasks. If the handwriting was correctly recognized, the error patterns were correctly identified in almost all cases, and equally accurate recommendations for the students' further development were given. This could significantly relieve teachers in the future. However, other mathematical areas are not sufficiently researched, and data protection issues remain unresolved.

\*\*Conclusion\*\*

We have seen that ChatGPT can correctly apply even complex mathematical methods. However, its independent problem-solving ability is in most cases limited to simple scenarios. The solutions often sounded professional and well-thought-out, although they were frequently at best incomplete or even incorrect.

The tendency to use GPT models for solving even moderately complex problems could undermine the development of critical thinking and independent problem-solving skills among students.

Consequently, everything suggests that while GPT can help students with simple tasks, its limitations in dealing with complex problems require careful integration into educational contexts. If tasks deviate from standard problems or increase in difficulty, all GPT versions quickly exhibit difficulties. This is particularly true for word problems that require a certain level of background understanding. Regardless of the focus on mathematics, the so-called prompt engineering always plays a crucial role. This refers to the user's ability to design inputs (prompts) in such a way that the model delivers the best possible answers.

GPT or comparable AI tools should therefore be considered as supplementary pedagogical tools and not as primary problem-solving aids. This limitation highlights the challenge of using AI in education: the inability of current models to fully grasp contextual nuances that are crucial for complex problem-solving and critical thinking. This also means that they should support but not replace traditional teaching methods and human guidance in advanced mathematical contexts.

Ultimately, the mathematical problem-solving abilities of AI are limited. ChatGPT should only be used by individuals who understand the mathematical background and is still unsuitable as a "blind" problem-solving or teaching tool in most cases. However, predicting how long this will remain the case is difficult. AlphaGeometry (not publicly accessible), in contrast to other LLMs, seems capable of solving problems from the International Mathematical Olympiad. Also, Llama 3.1 (soon to be publicly accessible) brings new developments. Furthermore, according to statements by the developers, ChatGPT5 is supposed to possess intelligence at the PhD level (doctorate in Germany). What exactly this means remains to be seen. One thing is certain: in the education sector, we must be prepared to react flexibly and quickly to new developments.

---

\*\*Dr. Oliver Straser, University of Education Freiburg\*\*

\*\*Aileen Fahrländer, University of Education Freiburg\*\*

\*\*Prof. Dr. Katja Maaß, University of Education Freiburg\*\*

\*\*Dr. Carina Spreitzer, University of Klagenfurt\*\*

\*\*Prof. Dr. Stefan Zehetmeier, University of Klagenfurt\*\*

---

\*\*References\*\*

Helfrich-Schkarbanenko, A. (2023). \*Mathematics and ChatGPT\*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-68209-8

Jordan, A., Ross, N., Krauss, S., Baumert, J., Blum, W., Neubrand, M., Löwen, K., Brunner, M., & Kunter, M. (2006). \*Classification Scheme for Mathematics Tasks: Documentation of Task Categorization in the COACTIV Project\*. Materials from Educational Research: Vol. 81. Max Planck Institute for Educational Research.

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., ... Zoph, B. (2023, March 15). \*GPT-4 Technical Report\*. http://arxiv.org/pdf/2303.08774

Spreitzer, C., Straser, O., Zehetmeier, S., & Maaß, K. (2024). Mathematical Modelling Abilities of Artificial Intelligence Tools: The Case of ChatGPT. \*Education Sciences\*, 14(7), 698. https://doi.org/10.3390/educsci14070698

# Ist Chat GPT ein Mathe-Ass?

### Mathematikunterricht in Zeiten von KI

Im digitalen Zeitalter hat sich die Art und Weise, wie wir lehren und lernen, drastisch verändert. Künstliche Intelligenz (KI) nimmt einen immer größeren Platz in unserem Alltag ein und bietet neue Möglichkeiten, Wissen zu vermitteln und zu erwerben. Ein bemerkenswertes Beispiel dafür ist Chat GPT, ein leistungsfähiges Sprachmodell (**LLM**, Large Language Model), das von OpenAI entwickelt wurde und über künstliche neuronale Netze funktioniert, die darauf ausgelegt sind, menschenähnlichen Text zu verstehen und zu generieren. Doch welche Rolle kann und sollte Chat GPT im Mathematikunterricht spielen? Kann diese KI tatsächlich als Mathe-Ass gelten und Lehrkräfte unterstützen, den Unterricht effektiver, interessanter oder auch differenzierter zu gestalten?

In diesem Artikel möchten wir Lehrkräften einen Einblick in die Fähigkeiten und Grenzen von Chat GPT im Mathematikunterricht geben. Wir untersuchen, wie diese Technologie funktioniert und wie sie mathematische Aufgaben löst. Gleichzeitig werfen wir einen kritischen Blick darauf, welche Herausforderungen im Umgang mit KI im Klassenzimmer beachtet werden müssen. Kann Chat GPT ein wertvoller Unterstützer im Mathematikunterricht sein oder sollten traditionelle Lehrmethoden weiterhin bevorzugt werden?

Die aktuellen Versionen von Chat GPT, basierend auf der GPT-4-Architektur, bieten verschiedene Modelle, die sich in ihren Fähigkeiten und Einsatzbereichen[[1]](#footnote-2) unterscheiden:

**GPT-4o**: Diese Version ist besonders effizient und leistungsstark bei der Lösung komplexer mathematischer Probleme und wissenschaftlicher Berechnungen. Sie bietet schnelle und präzise Ergebnisse und ist ideal für anspruchsvolle Unterrichtssituationen.

**GPT-4**: Die Hauptversion bietet umfassendes Verständnis und hohe Genauigkeit in einer breiten Palette mathematischer Konzepte. Sie eignet sich für detaillierte Erklärungen und tiefgehende Analysen im Unterricht.

**GPT-4-mini**: Diese kompaktere Version ist ressourcenschonend und für den mobilen Einsatz oder Umgebungen mit begrenzten Rechenressourcen gedacht. Sie kann grundlegende bis mittelschwere mathematische Probleme lösen und ist ideal für flexible und leicht zugängliche Lösungen im Klassenzimmer.

**GPT-3.5**: Die Vorgängerversion GPT-3.5 bietet ebenfalls starke mathematische Fähigkeiten, allerdings mit etwas geringerer Genauigkeit und weniger Kontextverständnis im Vergleich zu GPT-4. Sie ist dennoch eine solide Wahl für viele Unterrichtsszenarien und bietet eine kostenlose Alternative zu den neuesten Modellen (Stand: 08/2024).

Neben einfachen Texteingaben kann GPT auch Dateitypen verarbeiten, beispielsweise Texte analysieren und zusammenfassen. Die Antwort wird ebenfalls meist in Textform, aber auch in Dateien (z.B. Office Formaten) ausgegeben. Die kostenpflichtigen Varianten bieten darüber hinaus die Möglichkeit, größere Datenmengen auf einmal zu bearbeiten und Bilder erstellen zu lassen. Zudem lässt sich GPT4o für persönliche Bedürfnisse personalisieren, indem es beispielsweise bestimmte vorher eingespeiste Datengrundlagen besonders berücksichtigt. Darüber hinaus erhält man auch Zugriff auf andere veröffentlichte GPTs, einschließlich solcher, die mit Hilfe von Wolfram Alpha mathematische Probleme lösen.

### Wie funktioniert ChatGPT?

Zu Beginn zeichnete sich beim Einsatz von ChatGPT im mathemaischen Bereich ein ernüchterndes Bild ab. Bereits bei einfachen Anwendungen, wie zum Beispiel der Berechnung von Nullstellen von Funktionen, wurden die Lösungen in der Regel nicht korrekt berechnet. Stattdessen wurden sie entweder statistisch aus den Koeffizienten abgeleitet oder waren frei erfunden (halluziniert).

Inzwischen basiert die Lösung von Mathematikaufgaben durch Chat GPT auf einer mehrstufigen Verarbeitung von Text und mathematischen Symbolen. Der Prozess beginnt mit der **Tokenisierung**, einem entscheidenden Schritt, bei dem der eingegebene Text in kleinere Einheiten, sogenannte Token, zerlegt wird. Diese Token können Wörter, Satzzeichen oder mathematische Symbole sein. Sie werden durch das neuronale Netzwerk des Modells verarbeitet, das mit Millionen von Texten und mathematischen Aufgaben trainiert wurde. Das Modell erkennt Muster und wendet mathematische Regeln an, führt anschließend die mathematischen Operationen aus und generiert schrittweise eine Antwort. Es kann einfache Aufgaben direkt lösen und komplexere Probleme detailliert erklären. Bei der Anwendung komplexerer mathematischer Methoden oder Formeln greift GPT beispielsweise auf mathematische Bibliotheken in Python zurück, um die Aufgaben zu lösen. Dadurch wird klar, welche Schritte und Berechnungen durchgeführt wurden. Wie vorher angesprochen können spezialisierte GPTs auf andere Ressourcen zugreifen und erzielen damit andere Ergebnisse. Es existiert bereits Literatur, die aufzeigt, wie ChatGPT verwendet werden kann, um selbst komplexe mathematische Probleme zu lösen (Helfrich-Schkarbanenko, 2023). Dabei ist jedoch eine genaue Instruktion und die Wahl der mathematischen Hilfsmittel erforderlich.

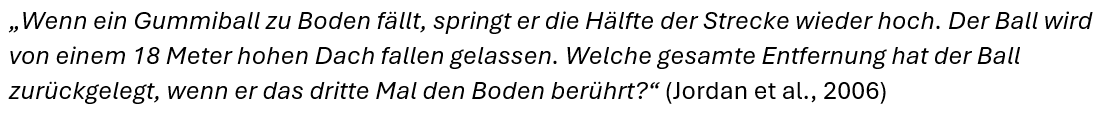
### Wie gut schlägt sich GPT beim eigenständigen Lösen von Aufgaben?

Dieser Frage ging der Entwickler OpenAI zunächst selbst nach und ließ GPT den Mathematikteil des amerikanischen Scholastic Assessment Tests (SAT), einen Teil des amerikanischen Hochschulzugangstests, absolvieren. Dabei schnitt GPT-4 im 89. Perzentil ab, also deutlich besser als der Durchschnitt. Zum Vergleich: Harvard-Studierende befinden sich in der Regel mindestens im 90. Perzentil. Allerdings müssen die Ergebnisse relativiert werden. Zum einen ist nicht klar, inwieweit diese Tests bereits zum Trainieren von GPT verwendet wurden und zum anderen enthält der SAT vergleichsweise simple, meist rechenlastige Aufgaben.

Andere Studien kamen zu differenzierteren Ergebnissen: Die Leistung hängt vom mathematischen Niveau, vom Gebiet und von der Art der Aufgabe ab. GPT ist gut dazu in der Lage, einfache Standardaufgaben zu lösen, bekommt aber mit höherer Klassenstufe zunehmend Schwierigkeiten. Abgesehen vom Schwierigkeitsgrad spielt bei Anwendungsaufgaben der Kontext und insbesondere die Formulierung eine schwerwiegende Rolle.

### Erstes Beispiel (Spreitzer et al., 2024)

Wir stellen ChatGPT einmal folgende Aufgabe und analysieren seine Lösung:

**

Ein Bild, das Text, Screenshot, Schrift enthält.

Automatisch generierte Beschreibung

Ein Bild, das Text, Screenshot, Schrift, Dokument enthält.

Automatisch generierte Beschreibung

Abbildung : Lösung von ChatGPT 3.5

Der Lösungsansatz mit der geometrischen Reihe (und das Ergebnis auch) ist falsch. Es liegt die Vermutung nahe, dass die Formulierung „…*springt er die Hälfte der Strecke …*“, GPT triggerte, die geometrische Reihe einzusetzen, da das Halbieren von Strecken häufig bei Anwendungsaufgaben zur geometrischen Reihe verlangt wird. Bei den neuen Versionen von ChatGPT taucht dieses Problem allerdings nicht auf.

### Zweites Beispiel: Tankaufgabe nach Blum

*Ein Bild, das Text, Schrift, weiß, Algebra enthält.

Automatisch generierte Beschreibung*Ein Bild, das Text, Screenshot, Schrift, Dokument enthält.

Automatisch generierte Beschreibung

Ein Bild, das Text, Screenshot, Schrift, Dokument enthält.

Automatisch generierte Beschreibung

Abbildung : Lösung der Tankaufgabe von ChatGPT 4

Die Lösung ist rein mathematisch betrachtet korrekt, fehlende Annahmen wurden korrekt identifiziert. Allerdings handelt es sich hier um eine rein mathematische Lösung, andere Faktoren wie ökologische oder zeitliche Faktoren fließen in die Argumentation auch auf Nachfrage nicht ein.

### Drittes Beispiel: Stauaufgabe nach Maaß

*Ein Bild, das Text, Schrift, weiß, Algebra enthält.

Automatisch generierte Beschreibung*

Ein Bild, das Text, Screenshot, Schrift, Dokument enthält.

Automatisch generierte Beschreibung

Ein Bild, das Text, Screenshot, Schrift, Dokument enthält.

Automatisch generierte Beschreibung

Abbildung : Lösung von ChatGPT 4

Die Lösung ist per se nicht falsch. GPT hat wie in der vorherigen Aufgabe auch fehlende Informationen identifiziert und nachvollziehbare Annahmen gemacht. Allerdings wird in der Lösung angenommen, dass die Autobahn nur eine Spur besitzt.

### Kann ChatGPT gut Mathematik lehren?

Die Heterogenität gilt in vielen Schulklassen als eine der wesentlichen Herausforderungen im Schulalltag und würde der Theorie nach in individuellen Betreuungssituationen am besten gemeistert werden. Die Vorstellung, dass Schüler\*innen von ChatGPT Aufgaben gestellt bekommen, GPT diese korrigiert und anhand der Lösungen individuelle Lernwege erstellt und betreut, klingt angesichts der oben genannten Defizite utopisch. Allerdings ist dies noch nicht endgültig erforscht und tatsächlich scheinen sich hier neue, zumindest sequenzielle Anwendungsfelder zu eröffnen. Im frühen Sekundarbereich werden beispielsweise rationale und Dezimalzahlen eingeführt. Gewisse Fehler bei Rechnungen, sogenannte systematische Fehler wie z.B. , lassen bei wiederholtem Auftreten auf Verständnisprobleme schließen. GPT ist ab Version 4 bereits recht gut dazu in der Lage, bei eingescannten Schüler\*innenlösungen derartige Fehlermuster zu identifizieren und entsprechende Vorschläge für Aufgaben zur Behebung zu liefern. Wenn die Handschrift korrekt erkannt wurde, wurden in fast allen Fällen die Fehlermuster richtig identifiziert und im gleichen Maße korrekte Empfehlungen für die Weiterentwicklung der Schüler\*innen gegeben. Dies könnte Lehrkräfte in Zukunft deutlich entlasten. Allerdings sind andere mathematische Bereiche nicht hinreichend untersucht und datenschutzrechtliche Fragestellungen nicht geklärt.

## Schlussfolgerung

Wir haben gesehen, dass ChatGPT selbst komplexe mathematische Methoden korrekt anwenden kann. Jedoch ist dessen eigenständige Problemlösefähigkeit in den meisten Fällen auf einfache Sachverhalte beschränkt. Die Lösungen klangen oft professionell und gut durchdacht, obwohl sie häufig bestenfalls unvollständig oder sogar falsch waren.

Die Tendenz, GPT-Modelle zur Lösung auch moderat komplexer Probleme zu verwenden, könnte die Entwicklung des kritischen Denkens und der unabhängigen Problemlösungsfähigkeiten der Schüler\*innen untergraben.

Folglich deutet alles darauf hin, dass GPT zwar Schülern\*innen bei der Lösung einfacher Aufgaben helfen kann, seine Einschränkungen im Umgang mit komplexen Problemen jedoch eine sorgfältige Integration in Bildungskontexte erfordern. Weichen die Aufgaben von Standardaufgaben ab oder nehmen an Schwierigkeitsgrad zu, so zeigen sich bei allen GPT-Versionen schnell Schwierigkeiten. Dies gilt insbesondere bei Textaufgaben, die ein gewisses Hintergrundverständnis voraussetzen. Unabhängig vom Fokus auf Mathematik spielt immer auch das sogenannte **Prompt Engineering** eine entscheidende Rolle. Dabei handelt es sich um die Fähigkeit des Nutzers, Eingaben (Prompts) so zu gestalten, dass das Modell die bestmöglichen Antworten liefert.

GPT oder vergleichbare KI-Tools sollten daher als ergänzende pädagogische Werkzeuge und nicht als primäre Problemlösungshilfsmittel betrachtet werden. Diese Einschränkung verdeutlicht die Herausforderung der KI-Nutzung in der Bildung: Die Unfähigkeit aktueller Modelle, kontextuelle Feinheiten vollständig zu erfassen, die für komplexe Problemlösungen und kritisches Denken entscheidend sind. Dies bedeutet auch, dass sie traditionelle Lehrmethoden und menschliche Lernbegleitung in fortgeschrittenen mathematischen Kontexten unterstützen, aber nicht ersetzen sollten.

Letzten Endes sind die mathematischen Problemlösefähigkeiten von KI begrenzt. ChatGPT sollte nur von Personen eingesetzt werden, die die mathematischen Hintergründe verstehen und ist als „blind“ einsetzbares Problemlöse- oder Unterrichtstool in den meisten Fällen noch ungeeignet. Doch ob oder wie lange das so bleibt ist schwer abzusehen. **AlphaGeometry** (nicht öffentlich zugänglich) scheint im Gegensatz zu anderen LLMs in der Lage zu sein, Aufgaben der Internationalen Mathematikolympiade zu lösen. Auch **Llama 3.1** (bald öffentlich zugänglich) bringt neue Entwicklungen. Zudem soll **ChatGPT5** laut Aussagen der Entwickler\*innen eine Intelligenz auf PhD-Niveau (Dr. in Deutschland) besitzen. Was dies genau bedeutet, wird die Zukunft zeigen. Sicher ist nur, dass wir im Bildungsbereich darauf vorbereitet sein müssen, flexibel und schnell auf neuere Entwicklungen zu reagieren.

*Dr. Oliver Straser, Pädagogische Hochschule Freiburg*

*Aileen Fahrländer, Pädagogische Hochschule Freiburg*

*Prof. Dr. Katja Maaß, Pädagogische Hochschule Freiburg*

*Dr. Carina Spreitzer, Universität Klagenfurt*

*Prof. Dr. Stefan Zehetmaier, Universität Klagenfurt*

Literaturverzeichnis

Helfrich-Schkarbanenko, A. (2023). *Mathematik und ChatGPT*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-68209-8

Jordan, A., Ross, N., Krauss, S., Baumert, J., Blum, W., Neubrand, M., Löwen, K., Brunner, M. & Kunter, M. (2006). *Klassifikationsschema für Mathematikaufgaben: Dokumentation der Aufgabenkategorisierung im COACTIV-Projekt*. *Materialien aus der Bildungsforschung: Bd. 81*. Max-Planck-Inst. für Bildungsforschung.

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., . . . Zoph, B. (2023, 15. März). *GPT-4 Technical Report*. http://arxiv.org/pdf/2303.08774

Spreitzer, C., Straser, O., Zehetmeier, S. & Maaß, K. (2024). Mathematical Modelling Abilities of Artificial Intelligence Tools: The Case of ChatGPT. *Education Sciences*, *14*(7), 698. https://doi.org/10.3390/educsci14070698

1. Wir gehen hier nicht auf die Möglichkeit ein, GPT per API zu verwenden. [↑](#footnote-ref-2)