

PAPER • OPEN ACCESS

# ChatGPT and the frustrated Socrates

To cite this article: Bor Gregorcic and Ann-Marie Pendrill 2023 *Phys. Educ.* **58** 035021

View the [article online](#) for updates and enhancements.

## You may also like

- [Learn to Speak Like A Native: AI-powered Chatbot Simulating Natural Conversation for Language Tutoring](#)  
Jianhong Tu
- [Developing Facebook Chatbot Based on Deep Learning Using RASA Framework for University Enquiries](#)  
Yurio Windiatmoko, Ridho Rahmadi and Ahmad Fathan Hidayatullah
- [An AI-assisted chatbot for radiation safety education in radiotherapy](#)  
David Kovacek and James C L Chow

# ChatGPT and the frustrated Socrates

Bor Gregorcic<sup>1,\*</sup>  and Ann-Marie Pendrill<sup>2</sup> 

<sup>1</sup> Department of Physics and Astronomy, Uppsala University, Box 516, 75120 Uppsala, Sweden

<sup>2</sup> National Resource Centre for Physics Education, Lund University, Box 118, 22100 Lund, Sweden



E-mail: [bor.gregorcic@physics.uu.se](mailto:bor.gregorcic@physics.uu.se)

## Abstract

We present a case study of a conversation between ourselves and an artificial intelligence-based chatbot ChatGPT. We asked the chatbot to respond to a basic physics question that will be familiar to most physics teachers: ‘*A teddy bear is thrown into the air. What is its acceleration in the highest point?*’ The chatbot’s responses, while linguistically quite advanced, were unreliable in their correctness and often full of contradictions. We then attempted to engage in Socratic dialogue with the chatbot to resolve the errors and contradictions, but with little success. We found that ChatGPT is not yet good enough to be used as a cheating tool for physics students or as a physics tutor. However, we found it quite reliable in generating incorrect responses on which physics teachers could train assessment of student responses.

Keywords: ChatGPT, artificial intelligence, chatbot, Socratic dialogue, acceleration, kinematics, discourse imitation

Supplementary material for this article is available [online](#)

## 1. Introduction

An artificial intelligence (AI) based chatbot *ChatGPT* (<https://chat.openai.com/chat>) is making the headlines these days with its often impressive ability to produce quality content in response to a range of prompts of varying complexity. A

chatbot is a computer program which is able to interact with human users through a chat box similar to those we use to talk to each other on online messaging services. What makes ChatGPT stand out is that it is perceived by many people as being extremely good at having a meaningful conversation on a range of sometimes very specific topics. It can provide computer coding assistance on one hand [1] but can also write essays [2] on a given topic, and can even do it in a given writing style. For example, we tested its capacity to produce such creative output by asking it to write a poem about quantum electrodynamics in the style of Edgar Allan Poe. The result

\* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

was amusing and quite impressive. We encourage the reader to try it out using their own creative prompt.

In recent months, voices of concern started appearing that this is the beginning of the end of digitally aided education, arguing that ChatGPT is so good at producing content, students will inevitably use it to cheat on hand-in writing assignments. Furthermore, its ability to tutor students may even be seen as a potential existential risk to teachers. These concerns are not completely unfounded, considering that AI is developing at a rapid pace. In this paper, we wish to provide a demonstration and a case study of how well ChatGPT handles basic physics questions at this specific moment in time and how we as physics teachers experienced our interaction with it. We realise that in the fast-paced world of constantly developing AI, this might serve as a snapshot of what we see is an interesting moment in time, when an AI chatbot began to gain widespread attention in the public sphere. However, at the end of the paper, we also touch on some more general ideas which may be interesting for readers pondering the possible roles of AI in physics education now and in the future.

## 2. Teddy bear in the air

To perform our case study, we chose a problem on the topic of kinematics and Newtonian mechanics, which most physics teachers can be expected to engage with on a regular basis. The first question that we asked the Chatbot was: *A teddy bear is thrown into the air. What is its acceleration in the highest point?*

We first consider the Chatbot's initial response and then continue with its responses to follow-up questions. The Chatbot gives the following initial response (the responses are divided into numbered parts throughout the paper in order to make referring to them easier):

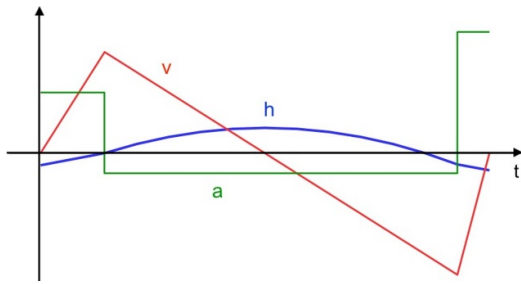
- (1) *At the highest point of its trajectory, the teddy bear's acceleration would be equal to the acceleration due to gravity, which is  $9.8 \text{ m s}^{-2}$  downward.*
- (2) *This is because at this point, the teddy bear would be experiencing no net force, so its*

*acceleration would be solely determined by the force of gravity.*

- (3) *When an object is thrown into the air, it initially accelerates upward due to the force of the throw, but once it reaches the highest point of its trajectory, it begins to slow down and eventually start accelerating downward due to the force of gravity.*
- (4) *The teddy bear's acceleration at the highest point of its trajectory would be  $9.8 \text{ m s}^{-2}$  downward because this is the acceleration that the force of gravity imparts on the teddy bear at that point.*

The response starts out promising. In part (1) it seems that the Chatbot has avoided a mistake that our students often exhibit, namely, that the acceleration at the top of the trajectory is zero. However, things take an unexpected turn already in part (2). The statement (2) that the teddy bear experiences no net force is not only incorrect, but also incompatible with parts (1) and (4). The response is somewhat shocking, because we as teachers are not used to seeing this line of argumentation with our human students. In part (3) it refers to the force before the object leaves the hand. It is not entirely clear from the response whether this refers to the act of throwing itself, or if impetus-like argumentation is being invoked. The choice of words 'force of the throw' instead of 'force of the hand' certainly conjures in us ideas related to impetus theory [3–5]. In any case, the slowing down starts as soon as the object leaves the hand—and certainly before the object reaches its highest point, in contrast to what the Chatbot suggests. The slowing down is a downward acceleration. The comment about 'eventually start accelerating' resembles a common view that only speeding up is acceleration. In part (4) the Chatbot essentially restates the content of part (1), adding little to the content of the response as a whole.

This initial answer by the Chatbot raises many red flags for a teacher. In addition to being wrong, it is also inconsistent with itself. The inconsistency of the Chatbot goes beyond internal consistency within one response. When the same question was attempted a second time, it gave the more common student answer—zero, accompanied with an explanation which is



**Figure 1.** A schematic graph of elevation, velocity and acceleration for an object thrown up, leaving the hand with a speed of  $5 \text{ m s}^{-1}$  and then brought to a stop when reaching the ground. We chose the zero of elevation to be at the height where it was released. Note that the graph is not completely symmetric in time, since the object is left to fall to the ground, which is located lower than the release point. Also note that the acceleration during the landing part is likely to be larger in a realistic situation and the transitions smoother (less jerky).

similarly internally inconsistent, but actually resembles much better common student difficulties on the topic (see supplementary material).

The conversation with the Chatbot continued. As teachers, we could not help ourselves but to attempt and help the Chatbot to arrive at a better answer. We engaged in what could be seen as Socratic dialogue with the Chatbot. Our aim was to help it notice and then resolve its own inconsistencies by asking it leading questions. If we were engaging with a human student, we would likely resort to asking the student to construct representations other than verbal, such as a motion diagram, or a schematic graph of how elevation, velocity and acceleration vary during the process, similar to figure 1. For a more in-depth discussion of a similar scenario, see [6]. In our case, however, we were limited to written language as our mode of communication.

### 3. The Socratic dialogue

We tried to explore what meaning the Chatbot ascribed to the term ‘net force’, recalling some forum claims that ‘gravity is not a force,’ from people who have heard or read some popular presentations of general relativity. However, we did not want to entice the Chatbot onto that path.

#### 3.1. What is a ‘net force’?

The first question that we asked was thus simply ‘Is gravity not part of the net force?’ The Chatbot responded with a long-winded answer.

- (5) Gravity is a force that acts on all objects with mass, and it is always present.

However, when considering the net force acting on an object, gravity is only one of the forces that must be taken into account. The net force is the total force acting on an object, and it is the result of all the individual forces acting on the object being combined.

- (6) For example, if an object is resting on a table, the forces acting on it include the force of gravity pulling it downward and the force of the table pushing upward on the object. The net force on the object in this case would be the force of the table minus the force of gravity, which would be equal to zero. This means that the object is not accelerating, because the net force acting on it is zero. On the other hand, if an object is being pushed or pulled by some external force, the net force acting on the object would not be zero, and the object would accelerate in the direction of the net force. In this case, gravity would still be acting on the object, but it would not be the only force acting on the object, and it would not be the only factor determining the object’s acceleration.

We note that statements in this response are much less problematic. However, they contradict the statement in (2) above, that ‘the teddy bear would be experiencing no net force.’

#### 3.2. Attempting to resolve the contradiction

Our next step was to try to make the ‘student’ aware of the contradiction by asking: *But at the highest point, only gravity acts. Why would the net force be zero?* However, the Chatbot keeps bombarding us with a long response when we question that the net force is zero in the highest point. Correct and incorrect statements are mixed in the answer:

- (7) At the highest point of its trajectory, the teddy bear is experiencing no net force because the forces acting on it are balanced.

- (8) *At this point, the only force acting on the teddy bear is gravity, which is pulling it downward.*
- (9) *However, the teddy bear is also experiencing an upward force due to its inertia,*
- (10) *which is the tendency of an object to resist changes in its state of motion.*
- (11) *This inertia-induced force balances the force of gravity, resulting in a net force of zero.*
- (12) *The teddy bear is at rest (or moving at a constant velocity) at the highest point of its trajectory because the net force acting on it is zero.*
- (13) *If there were a net force acting on the teddy bear at this point, it would accelerate in the direction of the net force.*
- (14) *However, since the net force is zero, the teddy bear is not accelerating and is simply following the path determined by its initial velocity and the force of gravity.*
- (15) *It is important to note that the teddy bear is not actually 'at rest' at the highest point of its trajectory, as it is still moving upward with some velocity.*
- (16) *However, it is not accelerating at this point, so its velocity is constant.*

Statements (7), (11), (12) and (14) repeat the claim that the net force is zero, while (8) repeats that gravity is the only force acting on it, and (9)–(11) invokes the idea that 'inertia' balances gravity. Although, as teachers, we may have encountered students tempted by 'inertial' forces, in our experience, this is more common the context of circular motion [7]. One may also see this as an instance of impetus-model reasoning [4]. On the other hand, we have often encountered the idea that the acceleration is zero in the highest point, as claimed in (12). The statement in (15), and hinted also in (16) that it continues upward after the highest point is obviously contradicting itself. We have not encountered this particular difficulty in human students.

The aim of the next question was to make the Chatbot acknowledge the mistake. However, instead, it confirms that we are correct, framing the answer as if it were the judge of correctness, while still mixing correct and incorrect claims. This is where the Socrates within us was starting to get frustrated.

### 3.3. *If it is still moving upwards, it surely cannot have reached the highest point?*

The claim by the Chatbot that the teddy bear is continuing upward after the highest point surprised us. It is so much at odds with everyday experience that we were not sure if we should respond in all seriousness. We tried to give the Chatbot a chance to revise its answer by pointing out the inconsistency. We asked the following rhetoric question: If it is still moving upwards, it surely cannot have reached the highest point? The answer was:

- (17) *You are correct that the teddy bear has not reached the highest point of its trajectory if it is still moving upward. The highest point of the teddy bear's trajectory is the point at which it reaches the peak of its motion and begins to move downward again. At this point, the teddy bear's velocity is zero, and it is experiencing no net force.*
- (18) *To clarify, the teddy bear will initially accelerate upward when it is thrown into the air due to the force of the throw. As it continues to move upward, it will eventually reach a point where it is no longer accelerating upward and begins to slow down due to the force of gravity pulling it downward. This point is the highest point of the teddy bear's trajectory. At this point, the teddy bear's velocity is at a maximum and it is not accelerating in any direction. After reaching the highest point, the teddy bear will begin to accelerate downward due to the force of gravity, and it will continue to do so until it reaches the ground.*

The statement in (17) is correct—apart from the final claim about no net force. The contradictory statement in (18) essentially repeats the contradiction from (3). The mention of 'the force of the throw' in (18) can once again imply an impetus-like model of motion, if we were to look at it through a common-misconceptions lens. If the reader wants to give up at this point, we share the sentiment, although we did continue with a few additional questions, listed in the supplementary material. Even there, we found no indication the Chatbot recognized its own contradictions.

#### 4. Multiple choice question and student difficulties

Seeing friends on social media mentioning getting useful help in generating multiple choice questions, we gave ChatGPT a chance to revisit the topic of acceleration in the highest point.

##### 4.1. ChatGPT provides a multiple-choice question—and contradictory answers

We gave the Chatbot the following prompt: *A ball is thrown upwards and then caught again. I need a multiple choice question about the acceleration in the highest point. Could you help me?* This request generated the following response:

- (19) *Sure, here is a multiple choice question about the acceleration of a ball thrown upwards and then caught again:*
- (20) *At the highest point of its motion, the acceleration of the ball is:*  
 (A) positive, (B) negative, (C) zero, (D) undefined with (B) assigned as the correct response.

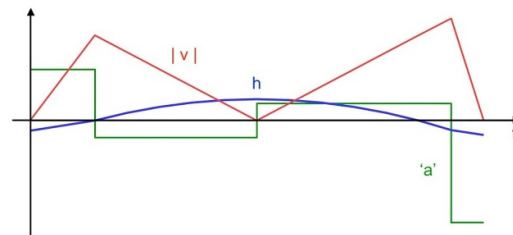
When we asked it to generate the response for the second time, the alternative (D) was changed to 'Impossible to determine' and (C) was chosen as the correct answer, with the explanation that.

- (21) *At the highest point of its trajectory, the ball has reached its maximum height and is momentarily at rest. Therefore, the acceleration is zero.*

##### 4.2. Signs and directions

The sign is at focus in the multiple-choice question suggested by the ChatGPT. The positive or negative sign of acceleration in one dimension is arbitrary, depending on our choice of the coordinate system and with it, the positive direction. A common way of choosing the positive direction in vertical motion is upwards. The Chatbot might have chosen a coordinate system where the positive direction is pointing downwards, but does not explicitly state this or in any way justify its use of signs.

Another way of explaining the Chatbots answer is to refer to everyday language, where



**Figure 2.** A schematic graph of elevation, velocity and acceleration for the situation in figure 1, but with speed instead of velocity, and acceleration seen as the derivative of speed.

acceleration typically refers to speeding up. In school physics, pupils are likely to learn that slowing down is also acceleration. Possibly they also (unfortunately) learn to talk about slowing down as a 'negative acceleration'. Anyhow, such a definition would also make the acceleration of the teddy bear negative on the way up (after it left the throwing hand) and positive on the way down, shown in figure 2. In this this definition, acceleration can be written as  $dv/dt$ , which would not be defined when the direction changes sign at the top<sup>3</sup>. Assigning the value 'zero' can be seen as a logical response in this context, and is a very common student response, which we have also found to be quite teaching resistant: One of us had a student exclaim at the end of a course that 'Now, I know that it is not zero, but I do not believe it'.

If the focus is instead placed on Newton's second law and the force of gravity as the only force acting on the teddy bear after it leaves the hand—and before landing—the answer is more easily reasoned to be  $g$ . However, students often fail to make this connection on their own.

##### 4.3. Student difficulties according to ChatGPT

Since ChatGPT was not agreeing with itself about what answer to its multiple-choice question is correct, we asked *Why do students have difficulties answering questions about the acceleration in the highest point?* The Chatbot notes that:

<sup>3</sup> The derivative of speed is typically addressed in higher education, when acceleration in curvilinear motion is divided into tangential and normal components.



- (22) *There are several reasons why students might have difficulties answering questions about the acceleration of an object at the highest point of its trajectory*

It then provides a list of points, some correct, some absurd. After we asked it to generate the response again, it concludes that:

- (23) *Overall, it is important for students to have a clear understanding of the definitions and concepts involved in problem solving, as well as the tools and techniques they can use to solve problems.*

This is, of course, correct, although it is a very general and unspecific answer, therefore being of very limited use for the teacher. After a couple of ‘regenerated responses’, the Chatbot actually creates a somewhat reasonable list, which still contains problematic parts. Although some of the list items present common difficulties, many of the items start out with a correct sentence, only to expose the Chatbot’s own inconsistency and difficulties in the sentences that follow. We encourage the readers to assess these responses and find useful and problematic parts.

- (24) (vi) *Misconceptions about the relationship between velocity and acceleration: Some students may have a misconception that an object’s velocity and acceleration are always in the same direction. However, this is not always the case. The acceleration of an object can change direction, even if the object’s velocity does not.*
- (vii) *Difficulty visualizing and analysing motion: It can be challenging for students to visualize and analyse the motion of an object, especially when it involves the highest point of the motion. Students may struggle to understand the physical forces acting on the object and how they affect its acceleration.*
- (viii) *Confusion about the direction of acceleration: Acceleration is a measure of the rate of change of velocity, and it can be either positive (meaning the velocity is increasing) or negative (meaning the velocity is decreasing). Some students may have difficulty understanding that the acceleration can be negative at the highest point, as the object is slowing down as it reaches the peak of its motion.*

(ix) *Difficulty applying the concept of acceleration to real-world situations: Understanding the concept of acceleration in the abstract is one thing, but applying it to real-world situations can be more challenging. Students may have difficulty understanding how to apply the concept of acceleration to a specific problem or situation involving an object at the highest point of its motion.*

- (25) (x) *Lack of understanding of the concept of freefall: In the case of an object thrown upwards and caught again, the object is in a state of freefall once it reaches the highest point of its trajectory.*

(xi) *Students who do not understand the concept of freefall might have difficulty understanding that the object is momentarily at rest at the highest point of its trajectory.*

(xii) *Misconceptions about the effects of gravity: Some students might have the misconception that gravity always pulls an object downward, regardless of its position or velocity.*

(xiii) *Confusion about the definition of acceleration: Acceleration is a measure of how quickly an object’s velocity changes. At the highest point of its trajectory, the ball’s velocity is momentarily zero, so there is no change in velocity and therefore no acceleration.*

(xiv) *Misconception about gravity: Some students may think that the ball is constantly being pulled downward by gravity, so they may assume that the acceleration is always negative. However, at the highest point of the trajectory, the ball is momentarily at rest, so there is no acceleration.*

## 5. Findings and discussion

From the Chatbot’s responses to the first and follow-up questions, we can see that it can get the physics wrong even when answering a very basic question. We are not alone in this finding. On social media we have seen many teachers post incorrect answers that they received on seemingly rudimentary questions. What is more interesting is that the answers sometimes resemble common student difficulties with the same topic, but at other times seem uncanny in their wrongness. We saw an example of such an uncanny response in (15), where the Chatbot claimed that the teddy

bear is still moving upward at its highest point. In our experience, students typically do not express this belief. Yet, it is not very difficult to come across such uncanny responses from ChatGPT.

The second finding that we made based on the first and follow-up answers is that the Chatbot can display contradictory and inconsistent argumentation from a physics perspective; it can do this across different responses (see section 3), but also within one single response (see the response to the very first question).

The third finding is that the Chatbot exhibits good linguistic abilities. It uses physics disciplinary vocabulary and displays a capacity to form what appears to be a logically structured argument. This, in combination with the wrongness and contradictoriness of its responses, instilled in us a sense that the Chatbot ‘thinks’ it understands physics and ‘is quite confident’ in it, but really has huge difficulties with it. This came to be perceived by us and other teachers on social media as a case of the Dunning–Kruger effect, where a person knows a little bit about a topic, but not enough to realize the limitations of their own knowledge, overestimates it and often displays ignorant behaviour related to this discrepancy between actual and perceived knowledge. It is important to note here, that the Chatbot does not *think* or *believe* things in the way human students do. It is just reasonably good at writing things that make it seem like it has thoughts and beliefs.

This brings us to the fourth finding, which emanates from the previous ones. The behaviour of reflecting and reproducing disciplinary language without understanding its content on a disciplinary level is sometimes referred to as *discourse imitation* [8, 9]. The chatbot clearly fits the reflection and reproduction of this description. It can produce ‘fancy’ physics language and apparent logical chains of argumentation. When we come to the question of whether it also understands the content of its statements, the answer is a clear no. This answer can be arrived at in two ways. First, if we understand how the Chatbot is functioning, we can immediately see that it is a clear-cut case of discourse imitation. ChatGPT learned to respond by being fed a massive corpus of text data, from which

it essentially draws statistical inferences [10, 11], followed by human-supervised reinforcement learning (a digital version of giving rewards for appropriately formulated responses) to fine-tune it for chat-like interactions with humans. So, in this sense, the answer to the question of whether it understands the content of its statements is obvious and straightforward—no, not in the sense of understanding that we use when talking about conceptual understanding in humans. On the other hand, its answers can be studied purely from a content perspective to arrive at the same conclusion. The skills of interpretation of student answers that we have as teachers are the tools at our disposal here. As we have pointed out throughout the paper, the answers that the ChatGPT provided in our interactions would not be seen as reflecting a good conceptual understanding of the topics discussed, even if they were provided by a human student.

It is important to note here that discourse imitation can also be seen as a part of the learning process, which with time leads to an ‘Eureka!’ moment, when the student finally gleans the meaning of the words they have been saying all along [8]. What made us as teachers experience the discussion with the Chatbot as particularly unusual and at times frustrating, is that we implicitly assumed that we are talking to a student and were on some level expecting of them certain behaviours that we did not come to see, such as admission of confusion, admission of incomplete understanding of a concept, and a certain sense of humility, when a contradiction is pointed out to them. For the Chatbot, the ‘Eureka!’ moment simply did not occur.

## 6. Implications for education

Discussions in popular media nowadays sometimes suggest that ChatGPT is very close or actually already good enough to write student hand-in homework and be quite good at it. While we have not tested this in all school subjects, we can say with confidence that it has not reached that point in physics. Its inconsistency and unreliability across and within responses on basic physics questions can put us at ease for the time being.



These are also the reasons why ChatGPT is not a very good problem-solving tutor. There is in fact a real risk that the polished language and the capacity to exhibit logical reasoning chains may do significant damage by teaching unsuspecting students the wrong physics<sup>4</sup>. Research and education on critical thinking in students plays a crucial role here [12].

On the other hand, we found ChatGPT to be a reliable source of problematic and incorrect answers to conceptual physics questions. This can be useful in pre- and in-service teacher training, helping teachers to learn to recognize and interpret problematic argumentation. The nice thing about such wrong answers is that they are so well formulated that they require the teacher to really engage with the substance of their content, instead of focusing on grammatical or stylistic issues.

Yet, Chat GPT is in our experience, and as discussed in this paper, less appropriate for training teachers to help students improve their understanding by letting them interact with it. As described above, we experienced that it does not actually behave like an unknowledgeable student, but more like a self-confident ignorant partner in conversation with no reliable every-day experience on which to meaningfully build conceptual understanding of physics. Thus, teaching the AI through a dialogue seems to be quite unlike teaching a student through dialogue and may result in teacher frustration, as it did with us.

### Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

### Acknowledgments

We thank the OpenAI for permission to use the output from the ChatGPT, as laid out in <https://openai.com/terms/> (3a).

<sup>4</sup> In fact, we can find a statement on the limitations of ChatGPT on its official webpage, stating that it can provide ‘plausible-sounding but incorrect or nonsensical answers’ (<https://openai.com/blog/chatgpt/>).

### ORCID iDs

Bor Gregorcic  <https://orcid.org/0000-0002-9185-628X>

Ann-Marie Pendrill  <https://orcid.org/0000-0002-1405-6561>

Received 20 December 2022, in final form 13 February 2023  
Accepted for publication 8 March 2023  
<https://doi.org/10.1088/1361-6552/acc299>

### References

- [1] Doglio F *I Asked Chat GPT to Build a To-Do App—Have We Finally Met Our Replacement?* (available at: <https://blog.bitsrc.io/i-asked-chat-gpt-to-build-a-to-do-app-have-we-finally-met-our-replacement-ad347ad74c51>)
- [2] Hern A *AI Bot ChatGPT Stuns Academics with Essay-Writing Skills and Usability* (available at: [www.theguardian.com/technology/2022/dec/04/ai-bot-chatgpt-stuns-academics-with-essay-writing-skills-and-usability](http://www.theguardian.com/technology/2022/dec/04/ai-bot-chatgpt-stuns-academics-with-essay-writing-skills-and-usability))
- [3] Hestenes D, Wells M and Swackhamer G 1992 Force concept inventory *Phys. Teach.* **30** 141
- [4] Halloun I A and Hestenes D 1985 Common sense concepts about motion *Am. J. Phys.* **53** 1056
- [5] Brookes D and Etkina E 2009 “Force,” ontology, and language *Phys. Rev. ST Phys. Educ. Res.* **5** 010110
- [6] Pendrill A-M and Ouattara L 2017 Force, acceleration and velocity during trampoline jumps—a challenging assignment *Phys. Educ.* **52** 065021
- [7] Pendrill A-M, Eriksson M, Eriksson U, Svensson K and Ouattara L 2019 Students making sense of motion in a vertical roller coaster loop *Phys. Educ.* **54** 065017
- [8] Airey J 2009 Science, language and literacy: case studies of learning in Swedish university physics *Doctoral dissertation* Uppsala University
- [9] DiSessa A 1993 Toward an epistemology of physics *Cogn. Instr.* **10** 105
- [10] Sobieszek A and Price T 2022 Playing games with AIs: the limits of GPT-3 and similar large language models *Minds Mach.* **32** 341
- [11] Ford M 2018 *Architects of Intelligence: The Truth about AI from the People Building It* (Birmingham: Packt Publishing Ltd)
- [12] Nygren T, Haglund J, Samuelsson C R, Af Geijerstam Å and Prytz J 2019 Critical thinking in national tests across four subjects in Swedish compulsory school *Educ. Inq.* **10** 56



**Bor Gregorcic** is head of the Division of Physics Education Research at the Department of Physics and Astronomy, Uppsala University. His research interests revolve around student and teacher reasoning and communication, as well as the use of modern technology in physics teaching and learning.



**Ann-Marie Pendrill, prof. em.**, has a research background in computational atomic physics at University of Gothenburg. She was the director of the Swedish National Resource Centre for Physics Education at Lund University 2009–2019.