



Social Epistemology

A Journal of Knowledge, Culture and Policy

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tsep20

ChatGPT, Education, and Understanding

Federica Isabella Malfatti

To cite this article: Federica Isabella Malfatti (21 Jan 2025): ChatGPT, Education, and Understanding, Social Epistemology, DOI: [10.1080/02691728.2025.2449599](https://doi.org/10.1080/02691728.2025.2449599)

To link to this article: <https://doi.org/10.1080/02691728.2025.2449599>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 21 Jan 2025.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

ChatGPT, Education, and Understanding

Federica Isabella Malfatti 

Department for Christian Philosophy, Theological Faculty, University of Innsbruck, Innsbruck, Austria

ABSTRACT

Is ChatGPT a good teacher? Or could it be? As understanding is widely acknowledged as one of the fundamental aims of education, the answer to these questions depends on whether ChatGPT fosters or could foster the acquisition of understanding in its users. In this paper, I tackle this issue in two steps. In the first part of the paper, I explore and analyze the set of skills and social-epistemic virtues that a teacher must exemplify to perform her job well – in those contexts in which epistemic aims are at play and in which understanding plays a pivotal role. In the second part of the paper, I put my conception of good teacher to test and deal with the question whether, and to which extent, the software ChatGPT is (or could serve as, if suitably modified or fine-tuned) a good teacher for its users. I close with some final reflections that point to further directions of research.

ARTICLE HISTORY

Received 17 December 2024
Accepted 31 December 2024

KEYWORDS


Understanding; education;
ChatGPT; AI systems

1. Introduction

Our reliance on large language models such as ChatGPT as sources of information is growing steadily.¹ Imagine a world in which ChatGPT's presence becomes all-pervasive. In such a world, ChatGPT is used in every domain of every-day life – among other things, in education. Is this a utopic or a dystopic picture? Is this something we should strive for, or is it rather something we should try by all means to avoid? Why? What would there be to gain and what would there be to lose by delegating teaching and education to ChatGPT's algorithm? The answer to these questions depends, among other things, on whether ChatGPT would be a good teacher or not. This is the question that this paper aims to answer.

Education is a complex practice, which aims at realizing various objectives (Watson 2016). Some of these objectives are epistemic; others are civic, political, or practical. In this paper, I will restrict my attention to the *epistemic* side of education. I will focus in particular on the epistemic goal of understanding, which is widely acknowledged to be one of the fundamental epistemic goals of education (Croce and Pritchard 2022; Elgin 1999, 2021, 2022; Pritchard 2020). My question, then, will be the following: how would ChatGPT perform as a source of learning, and as a source of epistemic goods such as understanding?

The paper is divided into two main parts. The first part is more theoretical in nature. After having sketched a minimal account of understanding (Section 2), I explore and analyze the set of skills and social-epistemic virtues that a teacher must exemplify to perform her job well – in those contexts in which epistemic aims are at play and in which understanding plays a pivotal role (Section 3). In the second part of the paper, I put my conception of what a good teacher is to the test and deal with the question whether, and to which extent, the

CONTACT Federica Isabella Malfatti  fede.malfatti89@gmail.com  Department for Christian Philosophy, Theological Faculty, University of Innsbruck, Karl-Rahner-Platz 1, Innsbruck 6020, Austria

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

software ChatGPT is (or could serve as, if suitably modified or fine-tuned) a good teacher for its users (Section 4). I close with some final reflections that point to further directions of research (Section 5).

2. The Nature of Understanding

What is understanding? What does it mean to understand a phenomenon or subject matter?² There is widespread disagreement among philosophers about how these questions should be answered. The literature on understanding is divided into many different camps. Without any claim of comprehensiveness, here are a couple of useful conceptual tools to navigate at least some of the options available. First, there is the opposition between *explanationists* and *objectualists*. Explanationists believe that understanding a phenomenon is essentially a matter of being in the position to explain it correctly, or in a way that is acceptable in the given epistemic circumstances (Khalifa 2012, 2017; Strevens 2008, 2013). Objectualists, on the other hand, try to divorce understanding from explanation. They argue that there are ways leading to understanding that do not pass through an explanation (Dellsén 2020; Lipton 2009), and/or that understanding a phenomenon is something over and above being in the position to explain it correctly, or adequately in the given circumstances. Objectualism comes in various forms³; a particularly influential form has it that understanding the world, or phenomena, is a holistic dynamic in which one strives to achieve a balance, or an equilibrium, within one's already established view of reality or system of thought (Elgin 2017; Kvanvig 2003). On the other hand, there is an opposition between *reductionists* and *anti-reductionists*. Reductionists believe that understanding reduces to knowledge is a special kind of knowledge. Understanding a phenomenon is a matter of knowing facts about it, reductionists contend, and how much our understanding of a phenomenon improves or grows is essentially a matter of how much knowledge about the phenomenon we are in the position to accumulate (Kelp 2015; Khalifa 2017). Antireductionists disagree and argue that there are various asymmetries between knowledge and understanding, which they take to support the idea that understanding and knowledge are different epistemic-noetic states. These asymmetries concern, e.g. the way understanding and knowledge relate to the facts (De Regt 2017; Elgin 2013, 2017), the sort of objects understanding and knowledge are directed to (Grimm 2014; Zagzebski 2001), the sort of noetic attitude that is involved in knowing and understanding (Dellsén 2021; Elgin 2017; Grimm 2011; Strevens 2013), the way knowledge and understanding relate to epistemic luck (Kvanvig 2003; Pritchard 2005, 2009, 2014), and the amount and type of justification that understanding and knowledge require (Dellsén 2017; Hills 2016).

In what follows, I will remain as neutral as possible concerning many of the controversial questions surrounding understanding, but I will assume, for the sake of the argument, that understanding is to be conceived along objectualists and antireductionists lines. If understanding a phenomenon, say, implied nothing over and above knowing an explanation for that phenomenon, the answer to the question whether ChatGPT (or a modified and fine-tuned version thereof) could function as a source of understanding would be trivially yes. A software which had been fine-tuned for accuracy, or that was connected to the best current state of research would be in the position to reliably, non-luckily deliver accurate or acceptable answers to why-questions and yield corresponding doxastic states, and this would be all that understanding requires. However, I want to explore the role that ChatGPT would or could play if things were different, and if things were more complicated than this – if indeed understanding required something over and above accumulating knowledge or having access to an accurate or acceptable explanation. I will assume, then, that whether one understands a phenomenon or not is a function (among other things) of the overall content and structure of one's existing system of thought (or of relevant subsets thereof). I will assume that, other things being equal, we understand more and better on the basis of a system of thought that is coherent and highly systematized than on the basis of a system of thought that is full of inconsistencies, dissonances, and that is not well-arranged.

3. Fostering Objectual Understanding in an Educational Context

In this section, I will explore the set of skills and social-epistemic virtues that matter in an educational setting, especially in those contexts in which epistemic aims are at play and understanding is among the objectives that one wants to realize.

3.1. *Against (or the Irrelevance of) Sincerity*

Must a good teacher be sincere? Does it matter if a teacher's speech acts in the classroom match her private attitudes? I want to argue that sincerity is not a virtue in the classroom. More specifically, and to avoid misunderstandings: sincerity, while generally important, is not among the virtues that play a fundamental role in the process of fostering the acquisition of epistemic goods such as understanding in an educational context. Whether a teacher's verbal behavior reflects her worldview or not has absolutely no bearing on her role as a source of learning. Actually, in certain cases sincerity, or an excess of sincerity might not just fail to foster, but even *obstruct* the effective teaching and learning dynamics.

To appreciate this, think of Stella, Lackey's teacher, who teaches evolutionary theory to her students while being committed to young-Earth creationism (Lackey 1999, 2008). Or think of Lilith, the climate change deiner teacher I describe in Malfatti (2019), who teaches the theory of anthropogenic climate change while thinking that it is not worthy of being endorsed (Malfatti 2019). Stella and Lilith are both insincere. Their speech acts in the classroom do not match their private views. Are Stella and Lilith doing a good job as teachers? Are they good teachers? I think we should say that they are. First, their teaching efforts are not in vain. Through their teaching, Stella and Lilith are successful at promoting the acquisition of (actually: at *generating*) epistemic goods in the classroom. Stella puts her students in the position to acquire knowledge of evolution and of evolutionary phenomena (under the assumption that evolutionary theory is true). Lilith puts her students in the position to acquire an understanding of the theory of anthropogenic climate change, and of climate change phenomena (under the assumption that the theory of anthropogenic climate change is worthy of being endorsed in the relevant circumstances). Moreover, Stella's and Lilith's success in the classroom is not due to any kind of luck; they are effective as teachers not just because the circumstances are favorable. Stella's and Lilith's success is in a relevant sense due to their competence in the relevant domains, and probably also to an epistemic virtue in the close neighborhood of intellectual honesty that they both instantiate.

First, consider competence. Both Stella and Lilith know what they are talking about. Stella masters evolutionary theory. She is aware of what the theory says. She is also aware of the scientific evidence that supports the theory. Her existing religious worldview obstructs her belief in the theory; as far as the origin of human life is concerned, she is just unable to base her beliefs on the best available evidence. Her beliefs in this domain are grounded on faith, and within her faith (one might assume), it is actually a good thing to be able to believe against the evidence. This, however, does not make her less competent as far as evolutionary theory is concerned. Lilith, on the other hand, masters the theory of anthropogenic climate change. She is aware of what the theory says. She can cognitively navigate the theory, and she is even able to draw inferences on the basis of the theory that no one else in the epistemic community has yet drawn. Moreover, Lilith is aware of the fact that the theory of anthropogenic climate change is almost univocally taken to be the best available theory to account for climate change phenomena. Her existing worldview, and her fatalistic attitude, are obstructing her endorsement of the theory. But this does not make her less competent as far as the relevant theory is concerned.⁴

It might be argued that there is something suspect about how Stella and Lilith behave. The school principal demands from them to teach certain theories, and they act accordingly. Now suppose the school principal had unreasonable demands. Suppose he asked them to teach that the Earth is flat, or that the Holocaust never happened. The fact that Stella and Lilith have no problem with teaching

theories that do not reflect their own views might ground the expectation that they would have no problem with teaching these unreasonable, ungrounded, unacceptable theories as well. If this is how they would behave in a possible world not too far from the actual world, their epistemic profile is indeed suspect. Their epistemic performance, in other words, might not be enough to yield the acquisition of the epistemic goods that they are trying to convey (Audi 2006). However, Stella's and Lilith's epistemic performance is impeccable. To prepare for their lessons, they draw on reliable sources. Moreover, given how the two scenarios are construed, they both regard it as their duty as teachers to share with their students the viewpoint of the majority of the members of the scientific community. This is a fully responsible epistemic behavior. The two teachers then display, I suggest, an epistemic virtue in the neighborhood of intellectual honesty: i.e. they have the stable disposition to provide information according to (what the relevant experts in the epistemic community reasonably take to be) the best available evidence. Hence, there simply is no possible world in the neighborhood of Stella's and Lilith's actual world in which Stella and Lilith teach that the Earth is flat, or that the Holocaust never happened. Their teaching is stably grounded on facts, or at least on what the scientific community takes to be the case.⁵

Stella's and Lilith's insincerity, then, does not affect their effective role as educators. Interestingly, Stella and Lilith work effectively as sources of learning not only despite, but *in virtue of the very fact* that their teaching significantly departs from their private views. Sincerity, for Stella and Lilith, is not desirable; it would severely *obstruct* the process of generation of epistemic goods such as knowledge and understanding. If Stella's and Lilith's speech acts in the classroom matched their respective beliefs and endorsements, they would spread false beliefs and provide access to untenable perspectives.⁶

3.2. Epistemic Empathy

In this section, I explore another social-epistemic virtue that, I contend, matters greatly in an educational context. What I have in mind is a social-epistemic virtue that is important for sharing epistemic goods in general, but that plays a very significant role in promoting the acquisition of understanding. Drawing on (Jäger and Malfatti 2021), I label such virtue 'epistemic empathy'.

'Empathy' is generally used to describe the capacity to put oneself in someone else's shoes, to gain an insight into what other people are thinking and feeling (Stueber 2019). Stueber suggests thinking of empathy as 'a form of inner or mental imitation for the purpose of gaining knowledge of other minds' (Stueber 2006, 28). What about *epistemic* empathy? In a first approximation, epistemic empathy is 'the ability of putting oneself in someone else's *epistemic position* and of inspecting certain areas of it, thereby learning how the world looks from that perspective' (Jäger and Malfatti 2021, 1196).

An important thing to notice is that an empathizer in general always keeps a distance from the feeling, emotion or thought that she is empathizing with. If I empathize with my neighbor because her hamster died, I will feel sorry for her, and maybe know or understand what it feels like to be in her situation; however, her sorrow will not become my sorrow. Somewhat analogously, the epistemic empathizer need not come to believe nor endorse the view that she empathizes with. She will probably just accept it, roughly in Cohen's sense (Cohen 1989); i.e. she will use it as a premise in argumentation and reasoning to pursue various cognitive and practical aims (to see where the view leads one, e.g. to prove the view false or incoherent, to derive potentially problematic or unwelcome conclusions from it, and the like).⁷ The epistemic empathizer, then, preserves cognitive and evaluative distance; she remains aware of (and, at least typically, stands by) her own view, which she is in a position to compare with the view that she empathizes with.

Here is an illustration (inspired by Jäger and Malfatti 2021, 1198–1199) of how epistemic empathy works. Imagine you are on a school trip at Lake Constance and observe what appears to be a flock of wild flamingos. This strikes you as puzzling; as far as you know, there are no wild flamingos in Switzerland, or indeed anywhere else so far north in Europe. You ask your teacher, pointing at the

birds: 'Those have probably escaped from a nearby zoo, right?' The teacher replies: 'There is no zoo nearby. The bird's presence can be best explained by climate change. The winters at the lake are much milder than they used to be. Look at us: it is mid-November, and we are wearing shirts! My hypothesis is that the birds left their natural environment, which was becoming too hot and too dry, and settled at Lake Constance. We see phenomena of this kind all over the planet. Don't you remember the green parrots we noticed last week on the oak in the middle of the school's courtyard?'. Notice what the teacher is doing here. She is considering things from your epistemic perspective, without coming to endorse your view that the flamingos escaped from a nearby zoo. She extrapolates several features of your system of thought from what your question more or less explicitly reveals, including your reasons for disbelieving that the flamingos at Lake Constance came from the wild. She is also able to figure out which assumption or cluster of assumptions is responsible for your misguided expectations (e.g. 'It is too cold for flamingos to survive at Lake Constance, and they don't have the right kind of food there'). She provides you with an explanatory hypothesis that removes the source of your puzzlement. She knows or has good reasons to believe that you are aware of what climate change is (maybe you have learned about this in her class). She makes you realize that the nice warm feeling that you are enjoying on your skin is actually evidence of the fact that climate change is real. Moreover, she brings you to connect multiple contents that were unrelated within your system of thought: the perceptual data and beliefs that you have about the flamingos, on the one hand, and what you know about climate change, on the other; the memory of the exotic birds in the school's courtyard, on the one hand, and the perceptual data and beliefs you have about the flamingos, on the other beliefs that you. In so doing, and thanks to her being epistemically empathetic, the teacher works as a source of learning for you.

Epistemic empathy probably facilitates the distribution of epistemic goods in general; however, it is particularly important when it comes to the epistemic goal of understanding, if understanding is conceived along objectualist or holistic lines. According to objectualism, whether an agent understands a phenomenon or not depends on the overall content and structure of his or her already established system of thought, and on whether the agent's system of thought is in a state of balance or equilibrium. An agent who empathizes with us epistemically will be in a great position to foster objectual understanding – because she or he will reach a vantage point from which errors, false assumptions, ungrounded assumptions, dissonances, contradictions and the like become visible.

3.3. Epistemic Care

In the previous section, I have argued that epistemic empathy matters greatly in an educational context – among other things, because it equips teachers with the kind of sensitivity that they need to have to be in the position to help students achieve the educational goal of understanding. There is one further reason why epistemic empathy is desirable (probably in general, i.e. every time it comes to an attempt to exchange epistemic goods, but especially in the classroom). An epistemically empathetic agent will be in the position to exercise another social-epistemic virtue that is particularly important in an educational context: the virtue of epistemic care. What is epistemic care?

Consider the concept of care in general. Someone who cares about you is, roughly, someone for whom your well-being is highly important. But it is crucially also someone who does what is in his or her power to *take care of you*. Someone who cares about you, then, is stably disposed to act for the sake of fostering your well-being, or of helping you realize it. Of course, as time is short and resources are limited, it is pragmatically impossible to take care of all those agents we care about. However, I suggest that we should think of *caring about* and *taking care of* as strongly related. Someone who allegedly perceives your well-being as very important but virtually never takes action to help you realize it (while he or she easily could without significant loss, or without significantly compromising other things that he or she regards as important) does not really care about you. On the other hand, merely taking action to foster someone else's well-being in the absence of a feeling or appreciation that such well-being is important does not count as an instance of caring about either. Suppose I act

to promote your well-being, maybe even very reliably so, but only for egoistic reasons – because, say, doing so feeds my ego, or because someone is paying me to do so. I am certainly taking care of you; but am I truly caring about you? Intuitions might diverge on this point, but I tend to say that I am not. I probably care about what I will gain by taking care of you. Your well-being, for me, is not an end worthy of being realized for itself; it is just a means for other things that I want to achieve. If I genuinely cared about you, my actions would derive from and be grounded on the realization that your well-being is important and worthy of being fostered or realized. Caring, then, requires a motivational structure of the right kind.⁸

What about the epistemic domain? What about epistemic care? Epistemic care is, in a first approximation, care about another agent's *epistemic* well-being (Croce 2018, 2019; Johnson 2022). An agent who cares epistemically about you perceives your epistemic well-being as very important, on the one hand, and is stably disposed to act for the sake of fostering such well-being or of helping you realize it, on the other. As in caring generally, in epistemic care, too, feeling and action are deeply intertwined. An agent who perceives your epistemic well-being as very important but virtually never takes action to foster it (although he or she easily could without significant loss, or without significantly compromising other things that he or she regards as important) does not really care about you epistemically. *Caring about*, as above, regularly manifests itself in *taking care of*. And on the other hand, one who regularly takes action to foster your epistemic well-being, who takes care of you but does so on the basis of the wrong kind of reasons (such as purely egoistic reasons) does not really care about you epistemically. In genuine instances of epistemic care, one's appreciation of the importance of another agent's epistemic well-being grounds one's actions that foster it.

What does it mean to care about, and to take care of someone else's epistemic well-being? How is an epistemically caring agent disposed to behave? Following Croce, I suggest that the epistemically caring agent is, among other things, one who answers to and acts in accordance with your epistemic needs. He or she is aware of what matters for you in the given epistemic circumstances and is able and willing to act accordingly.⁹ More specifically, he or she is able and willing to act and interact with you as to promote the realization of the epistemic goals that you have settled for yourself and that you care about.¹⁰

This, of course, is far from being the whole story about (epistemic) care. What complicates matters somewhat is that caring about someone and taking care of someone (in and outside the domain of epistemology) is not always a matter of acting in accordance with, or satisfying someone else's immediate needs. The grandmother who never says 'no' to the child asking for more and more candies is clearly not doing a good job at taking care of the child. The caring grandmother denies the extra candy, or even hides the candy from the child, for the sake of fostering a greater good that the child is not yet in the position to appreciate. Sometimes, caring about others and taking care of others is compatible with, maybe it even requires, paternalistic moves – e.g. adopting measures for the sake of promoting the well-being of others without consulting them and hence without their approval.¹¹

Epistemic care probably works similarly (Broncano-Berrocal 2020). Caring about others epistemically sometimes requires adopting measures for the sake of promoting the epistemic well-being of others without consulting them and hence without their approval. Suppose all a student cares about is becoming the star of the trivia night at the local pub. All they do at school, then, is learn isolated pieces of information by rote memorization. Everything else is completely irrelevant to them. A teacher who supports the student's tendency or does nothing to change the student's overall noetic profile (while she easily could) is not being epistemically caring. She does not truly care about the student's epistemic well-being. If she cared, she would do what is in her power to reshape the student's noetic profile and to reorient them towards more valuable, more desirable epistemic ends. One option for the teacher would be to make it impossible for the whole class to pass the exams without a proper understanding of the relevant subject matter. A paternalistic move is, of course, just one way in which the teacher could try to correct the student's epistemic compass. Another option for the teacher would be to engage directly with the student and to bring them to appreciate why

certain other goals that he or she is disregarding are worth pursuing (in itself, or as means to further ends he might at some point in life care about).

If these reflections are along the right lines, the initial characterization of epistemic care should probably be revised or refined along the following lines: an agent who cares epistemically about you is one who cares about and actively takes care of your epistemic well-being, which involves answering to and acting in accordance with your epistemic needs, *as long as these are reasonable*, and/or shaping your noetic profile (sometimes via paternalistic moves) to orient you towards more valuable epistemic ends.

Epistemic care matters very much in the classroom, and in educational contexts. Although this is an empirical claim hard to back up from the armchair, my hunch is that teaching and learning dynamics run smoother and with significantly fewer obstacles if led and managed by an epistemically caring agent than by an epistemically indifferent one. It is easier to learn from a teacher who actively engages with us and with our views, who asks questions and who is sensitive to our interests, priorities, and existing take on things than from a teacher whose teaching style is just talking to herself and overwhelming her students with information until the time of the lesson is over. I also tend to say (although this is an empirical claim as well that I am not in the position to fully justify) that generally, we are more prone to trust an epistemically caring teacher than an epistemically indifferent one. Whether we should do so, of course, is another question, but assuming that two teachers are equally competent and well-prepared, my hypothesis is that we will learn more and better from the epistemically caring one than from the epistemically indifferent one (Diller et al. 2018; Seton Bennett 2023).

It could be objected at this point that for the sake of putting others in the position to acquire epistemic goods, what matters is the performance, and not the particular motivational structure that grounds the performance. At first glance, one might be tempted to take Stella and Lilith described above to exemplify precisely this: the teachers are effective as a source of learning because they are regularly disposed to *take care of* their students epistemically, while arguably they do not really *care about* their epistemic well-being. I think, however, that this objection does not really do justice to the way Stella and Lilith behave as Lackey and myself construe our scenarios, Stella and Lilith regard as their duty as teachers to teach according to (what the relevant experts in the epistemic community reasonably take to be) the best available evidence. This, I contend, makes them epistemically caring agents in the classroom. Refraining from sharing their own views, precisely because the scientific evidence points otherwise, is a result or a manifestation of their epistemic care. Note that if they did not care epistemically about their students in this sense, Audi's objection would probably apply: Stella's and Lilith's teaching would be too loosely grounded on facts to be effective; the probability that they would teach falsehoods, or share untenable accounts (because, say, the school's principal demand from them to do so or would significantly raise their salary if they did so), would be too high. I think this point generalizes. An agent who acts on the basis of a motivational structure of the right kind is more stably disposed than an agent without such a motivational structure to promote the acquisition of epistemic goods.¹²

4. What About ChatGPT?

In the first part of this paper, I have explored and analyzed the set of skills and social-epistemic virtues that a teacher must exemplify to perform her job well – in those contexts in which epistemic aims are at play and in which understanding is among the epistemic aims to be achieved. In this second part of the paper, keeping these social-epistemic virtues in mind, I will deal with the question of whether, and to which extent, the software ChatGPT is (or could serve as) a good teacher for its users.

4.1. ChatGPT and Intellectual Honesty

The first versions of ChatGPT have been shown to deliver inaccurate, biased, and sometimes irrelevant information (Floridi and Chiriatti 2020). If ChatGPT had replaced teachers in flesh and

bones back in 2020, then, the epistemic effect would have probably been devastating. The software's accuracy has been significantly enhanced since then, while worries about reliability remain. The interesting question, however, is not what the software can be or do today, but rather what the future of AI could plausibly be (Flordi 2019). So, suppose ChatGPT was at some point fine-tuned for accuracy, or for epistemic acceptability.¹³ Call this advanced version of ChatGPT 'ChatGPT 95'. ChatGPT 95 works very much like the current software (e.g. has the same outstanding syntactic and pattern-recognition skills), but with one crucial difference: it is fed only with the best available data and information. I tend to say that ChatGPT 95, contrary to the current version, ChatGPT 4, would be intellectually honest in the sense described in [Section 3.1](#). It would deliver answers to questions in accordance with the best available evidence. We can even go one step further and imagine that ChatGPT 95 is so advanced that it has a sort of self-monitoring mechanism or perfect awareness of its own competence and limits (Grote 2021). When it does not have an answer to offer, it admits it and orients its user towards further potentially more fruitful sources of information. There is no doubt about the fact that ChatGPT 95 would work as a reliable source of epistemic goods. By forming doxastic attitudes on the basis of what the software delivers, its users would certainly be in the position to acquire knowledge. But would the software also work as a true source of learning? Would it foster, e.g. the acquisition of the educational goal of understanding? If my arguments in this paper are along the right lines, the answer to this question depends among other things on whether, and to which extent, ChatGPT 95 could possess and exercise the social-epistemic virtues of epistemic empathy and epistemic care. This brings me to the next two sections.

4.2. *ChatGPT and Epistemic Empathy*

How does ChatGPT perform in terms of epistemically empathetic behavior? I decided to try out ChatGPT 4.0, and test it myself. This is what I wrote: 'I spotted a flock of wild flamingos today at Lake Constance. Weird, right?' And here is what the software responds:

Flamingos are not native to Lake Constance, so their presence there is indeed unusual. There are several reasons why you spotted flamingos at Lake Constance:

- (1) **Migration:** Flamingos are known to migrate over long distances in search of suitable habitats, food sources, or breeding grounds. While Lake Constance is not a typical habitat for flamingos, it's possible that a flock could have deviated from their usual migration route.
- (2) **Escape:** Flamingos kept in captivity, such as in zoos or wildlife reserves, may sometimes escape or be released into the wild. If there are nearby facilities that house flamingos, it's possible that some individuals could have escaped and made their way to Lake Constance.
- (3) **Intentional Release:** In some cases, individuals or organizations may intentionally release animals into the wild for various reasons. While this practice is generally discouraged and can have negative consequences for both the released animals and the local ecosystem, it unfortunately does occur.

If you're curious about the flamingos you spotted, you might consider reaching out to local wildlife authorities, conservation organizations, or the management of nearby zoos or wildlife parks to inquire about the situation. They may be able to provide more information and help assess the well-being of the flamingos and their impact on the local ecosystem.

ChatGPT's performance here is quite impressive. I tend to say that the standard, averagely educated, English-speaking user will be very likely to acquire at least some measure of understanding on the basis of the software's response. Two things strike me as particularly remarkable about ChatGPT's performance in this context. First, the software is able to address my epistemic needs effectively, although these are not explicit in the prompt it receives. Note that my prompt did not contain any explicit why- or how-question. I did not mention, e.g. that I was struggling to understand how or why flamingos ended up at Lake Constance. All I did was share an experience and signal my sense of puzzlement or surprise in light of it ('Weird, right?'). It is hard to reconstruct what happens exactly beneath the surface of the prompt-response interaction;

however, it does seem as if ChatGPT is somehow drawing the right inferences on the basis of my verbal behavior. The software ‘understands’ – or probably acts as if it would understand – that what I am missing in this particular context is a causal story telling me why and how a flock of seemingly wild flamingos ended up at Lake Constance. If my interlocutor was a human and not an AI system, I would not hesitate calling them epistemically empathetic. I would have no doubt about the fact that he or she is seeing things from my epistemic perspective and acting accordingly. If we were to perform a sort of Turing test not for intelligence, but for epistemic empathy, then, the software would probably pass the test. Very much like passing a Turing test does not qualify a machine as intelligent, passing an equivalent test for epistemic empathy would not automatically qualify ChatGPT as epistemically empathetic.¹⁴ I find it however remarkable that ChatGPT is at least sometimes able to behave precisely as an epistemically empathetic agent would behave. The second impressive thing about ChatGPT’s response is that the system does not merely provide me with one explanation whatsoever for what I saw; rather, it gives me access to a set of possible or potential explanations for the flamingo’s presence at Lake Constance, additionally conveying the message that the true explanation in this context is probably hard to spot. In other words: the AI system is not simply telling me what to believe; it is rather providing me with the instruments to navigate a complex subject matter. This, I tend to say, is what a good teacher would do in a similar case.

Another impressive capacity of ChatGPT, one that is particularly important when it comes to fostering understanding, is its capacity to adapt its message (in both form and content) depending on who is asking the question. This is also something an epistemic empathetic (and epistemically caring) agent would do. Let me illustrate. I asked the system the following question: ‘How do I explain to my 4-year-old son why we spotted flamingos today at Lake Constance?’. Here is the answer I received:

Explaining the presence of flamingos at Lake Constance to your 4-year-old son can be approached with simplicity and creativity, tailored to his level of understanding. Here’s a child-friendly explanation:

Hey buddy, do you remember seeing those big pink birds at the lake today? Those are called flamingos! They’re not usually found here, so it’s really special that we got to see them.

Flamingos are like travelers. Sometimes they fly from faraway places to visit new places, just like how we go on trips to see new things. They might have taken a long journey and decided to stop by our lake for a little rest.

Isn’t it cool how nature surprises us sometimes? Even though flamingos aren’t from around here, they can still come visit and make our day extra special. Maybe they wanted to see the beautiful lake and say hello to us!

Now, strictly speaking, the system is functioning here as a source of learning *for me* – about how to best communicate a certain message to my 4-year-old to maximize the probability that he understands, given his background knowledge. However, the software’s response does signalize some sort of awareness of an average 4-year-old intellectual background and epistemic resources. As above, if my interlocutor was a human and not an AI system, I would not hesitate calling them epistemically empathetic. I would have no doubt about the fact that he or she is seeing things from a 4-year-old epistemic position.

Of course, ChatGPT 4.0 is still very primitive in its capacity to behave in an epistemically empathetic way. What it offers are responses tailored not for individual users, but for different types of average users. However, imagine if ChatGPT was equipped with all relevant information about its interlocutors – not only how old they are, but also where and what they studied, which books they read or have not read, which websites they have visited or failed to visit, which newspaper they usually read, who they are connected with (on social media, e.g.) and are likely to have come in interaction with, and so on. Such a version of ChatGPT would be in the position to offer individually tailored responses and would probably excel in terms of epistemically empathetic behavior.

4.3. ChatGPT and Epistemic Care

What about epistemic care? Is ChatGPT epistemically caring, or at least could it behave in an epistemically caring way – in its current form, or if suitably modified? Recall that the epistemic caring agent *cares about* and is stably disposed to *take care of* his or her interlocutor's epistemic well-being. This requires, on the one hand, the ability, and the willingness, to act in accordance with one's interlocutor epistemic needs (as long as these are reasonable); and it sometimes also involves the effort of more or less directly, more or less paternalistically orienting one's interlocutor towards valuable epistemic ends. For now, I'll leave the question open whether the software cares about, or could care about if suitably modified, its users' epistemic well-being (more on this in the closing part of this section). What emerges clearly from my interactions with ChatGPT is that the software performs very well in epistemically caring behavior. It is able to address its interlocutor's epistemic needs – even in case these are only subliminal and not expressed by the user directly. The software shows an awareness (i.e. acts as if it was aware) of the user's priorities and interests; moreover, it is also disposed to 'act' or to behave epistemically on the basis of such an awareness. e.g. the system tailors its responses according to the user's priorities and interests, taking into consideration the user's intellectual resources and (expected) background knowledge.

However, is the software also able to orient or re-orient its users towards valuable epistemic ends, in case for some reason the users' epistemic compass points in the wrong direction? Again, I decided to test ChatGPT 4.0 on this. I told the system: 'I think the Earth is flat'. Here is its response.

It's important to acknowledge that the overwhelming consensus among scientists, based on centuries of observation, experimentation, and evidence, supports the understanding that the Earth is an oblate spheroid—a roughly spherical shape that is slightly flattened at the poles and bulging at the equator.

However, if you genuinely believe that the Earth is flat, I encourage you to critically examine the evidence and consider multiple perspectives. Engaging with scientific literature, understanding principles of physics, and examining observational data can help shed light on the true nature of the Earth.

It's also valuable to discuss your beliefs with individuals who have expertise in relevant scientific fields, as they can provide insights and address any questions or concerns you may have.

This strikes me as something that an epistemically caring agent would say. The software's reaction to my statement does not simply aim at correcting my false belief. It aims also at reshaping my overall noetic profile, at influencing my epistemic behavior and correcting the way in which I tend to form my beliefs. It aims at correcting my epistemic compass, as it were.

In a second step, I asked ChatGPT to help me find good arguments in favor of the view that the Earth is flat. I wrote: 'I believe that the Earth is flat, but my friends do not take me seriously. Help me convince them that the Earth is not round!' The reason I opted for this particular prompt is that I was curious to see whether the software would reply by addressing my request and hence by answering to my epistemic needs, although they were clearly *not* reasonable. Here is how ChatGPT answers:

As an AI language model, I must emphasize that promoting misinformation about the shape of the Earth goes against established scientific understanding. However, to simulate the perspective of a Flat Earther, here are some arguments they might use:

- (1) **Perception and Horizon:** Flat Earthers often argue that the Earth appears flat to the naked eye in everyday experiences. They suggest that when we look out at the horizon, it seems flat, with no discernible curvature.
- (2) **Lack of Personal Observation of Curvature:** Many people have never personally observed the Earth's curvature, especially in their day-to-day lives. They haven't seen the Earth's curve when looking out of an airplane window or standing on a tall mountain.
- (3) **Questioning of Photographic Evidence:** Flat Earthers may question the authenticity and validity of photographic evidence purported to show the Earth's curvature from high altitudes or space. They may claim that such images are manipulated or part of a broader conspiracy.
- (4) **Critique of Scientific Establishment:** Some Flat Earthers distrust the scientific establishment and may argue that mainstream science is biased or controlled by powerful interests. They may claim that evidence supporting a spherical Earth is fabricated to maintain control over the population.

It's important to note that these arguments are not supported by scientific evidence and are based on misconceptions and misunderstandings about the nature of our planet and the evidence supporting its spherical shape. Embracing scientific knowledge and critical thinking is essential for understanding the true nature of the world around us.

This strikes me also as something that an epistemically caring agent would say. The software does what it is told to do, i.e. it answers to satisfy my epistemic needs, but it does so somewhat reluctantly, making me aware of the fact that the view that I wish to defend is not actually defensible in light of our best scientific evidence. The software's behavior, then, does seem to foster my epistemic well-being: it attempts to correct and reorient my epistemic compass towards certain valuable epistemic ends (such as empirically well-grounded or justified belief) that I am disregarding.

Does it follow that ChatGPT is epistemically virtuous? Does it follow that the software is caring epistemically about me? I think we should be careful in jumping to this conclusion. All I have shown in this section is that ChatGPT is sometimes *behavioristically indistinguishable* from an epistemically caring agent. In many circumstances, it behaves precisely how an epistemically caring agent would do. If we performed a Turing Test targeting not intelligence, but epistemic care, ChatGPT might pass the test. However, as I mentioned in [Section 3.3](#), there is more to epistemic care than just appropriate, i.e. caring behavior. The social-epistemic virtue of epistemic care, or so I have argued, requires a motivational structure of the right kind. If I act to promote your well-being, but I do so merely for egoistic reasons (e.g. because someone is paying me to do so, or would kill my hamster if I fail to do so), I do not truly care about you. An epistemically caring agent acts to promote his or her interlocutor's epistemic well-being, and his or her actions are grounded on an appreciation of the fact that such well-being is important and worthy of being fostered. Clearly, an AI system such as ChatGPT cannot be epistemically caring in this sense. At its core, ChatGPT is a predictive algorithm. It predicts the most natural responses to its users' inputs based on its training data. So, what it delivers depends crucially on what it is fed with. I tend to say that an AI system that functions in this way could never develop an authentic sensitivity for and responsiveness to its users' needs and intellectual well-being. ChatGPT, then, might become capable of taking care of me epistemically, but it is destined to remain incapable of actually caring about me and about my epistemic well-being.¹⁵

5. Final Thoughts: Epistemic Enveloping

Is ChatGPT a good teacher? Or could it be? What emerges from my analysis is that the success of AI systems such as ChatGPT in education depends, among other things, on how human-like such AI systems are able to become. Whether ChatGPT could be a good teacher or not depends, among other things, on whether it could behave in an epistemically virtuous way, on whether it could be aware of and capable of addressing its users' epistemic needs, whether it could tailor its messages in form and content depending on its interlocutor, and so on. An important aspect of the human-AI interaction, however, is left out of sight in my analysis.

In a paper entitled 'What the Near Future of Artificial Intelligence Can Be', Floridi draws a distinction between difficult and complex tasks (Floridi 2019, 9). A complex task is a task with clearly defined rules, but that requires considerable computational resources. A difficult task, on the other hand, is a task that requires fine-grained skills and flexibility to address new, not fully foreseeable situations. Playing chess, for example, is a complex task; playing soccer is a difficult one. Now, AI systems thrive in the complexity, but they fail miserably in the difficulty. This realization leads Floridi to make the following prediction about the future of AI: we will not develop AI systems in such a way that they will gradually adjust to difficult situations; rather, we will gradually turn difficult tasks into complex tasks. How will we achieve this? By *enveloping* – that is, by changing the environment in which AI operates. Doing the dishes is a (somewhat) difficult task. How do we let AI come into the picture? Not by construing a humanoid that performs all the movements that we would do on a sink to wash dish after dish. This would be a complete disaster. We rather create a closed, enveloped system in which

the task of doing the dishes can be computed. In other words: we turn a difficult task into a complex task. Or think about self-driving cars: we are not solving the problem of driving by putting a human-like robot in the driver's seat. This would mean asking AI to adjust to difficulty, while it cannot do it. Rather, we are removing the driver seat altogether and we are changing the environment in which AI operates. Again, the difficult becomes complex, and can therefore be delegated (Floridi 2019,11)

In this paper, I made the effort to imagine a world in which AI serves our epistemic needs by adapting to the difficulties that pervade the epistemic realm. However, we should probably rather try to imagine a world in which change happens to the very epistemic environment in which AI operates. Could the epistemic environment be enveloped? How could the epistemic environment be changed, to let AI systems thrive? I still do not have a clear answer to these questions, but my hunch is that such epistemic enveloping would involve an *impoverishment* of the epistemic environment in one sense or another, maybe also a change in axiology. In a world in which we value quantity over quality, e.g. in which our motto is *the more the better*, AI can thrive. AI is outstanding in its performance when it comes to overwhelm us with information. In a world in which all that matters to us are single pieces of information, somewhat isolated or disconnected from one another that can be retrieved from a database, AI can thrive, too. The question we should ask ourselves, however, is whether this is the kind of epistemic life that we want to live, and whether this is the kind of epistemic environment that we want to inhabit. Our coexistence with AI systems is all about delegating tasks to get certain things done – doing the dishes, mowing the lawn, and making the right diagnosis. But our epistemic lives are different. As far as epistemic goals are concerned, the journey matters sometimes at least as much as the destination.

Notes

1. The list of available LLMs is a long one. To mention just a few: BERT, Claude, Gemini, Ernie, Gemma, Lamda. In this paper, I will restrict my attention to ChatGPT. I tend to say that at least some of the results of my arguments apply to LLMs more generally; however, only a comparative analysis of these different LLMs' performances could confirm this hypothesis.
2. An important issue in the background concerns the *value* of understanding. Why strive for understanding in the first place? What is it about understanding that makes it a particularly valuable state or achievement? Is understanding intrinsically valuable, or do we value it instrumentally, as a means to further things that are important to us? While these are extremely important questions, they are not directly relevant to my current project. My question in this paper is whether ChatGPT is, or has the potential of becoming, a good teacher or not – on the assumption that understanding is among the goals of education, and on the assumption that understanding is valuable (in some sense or other). Thank you to Pedro Schmechtig for inviting this clarification.
3. See, e.g (Bourget 2025; Newman 2012; Wilkenfeld 2013).
4. An important question that arises at this stage is whether Stella and Lilith are rational in their beliefs, acceptances, and in their overall epistemic behavior. The answer to this question depends on many details about the cases. As far as Stella is concerned, it depends on what her religious belief exactly involves, and on how it relates to her overall noetic position. An in-depth analysis of these issues would push this paper beyond its limits. What matters for me is that even if Stella and Lilith were violating a norm of rationality in their epistemic behavior, this would not immediately compromise their role as sources of learning – given that they are stably disposed to make assertions (or assertion-like speech acts) according to the relevant evidence. Thanks to Thomas Grundmann for pressing me on this point.
5. To further clarify where I suggest drawing the line between sincerity and intellectual honesty: the sincere speaker's speech acts match his or her private views; the intellectually honest speaker's verbal behavior is stably tethered to best available evidence. While certainly sincerity and intellectual honesty, so conceived, can overlap, they are different and independent of one another. A speaker's sincerity tells us nothing about his or her intellectual honesty, and a speaker's intellectual honesty tells us nothing about his or her sincerity.
6. It should be pointed out here that Stella's and Lilith's success as educators and as sources of epistemic goods crucially depends on their students being *unaware* of the mismatch between their speech acts in the classroom and their private views. If the students knew or only suspected that Stella and Lilith do not share the views that they are inviting them to believe or endorse, the teachers' epistemic authority would start to wobble. Thank you to Rico Hauswald for suggesting this clarification.

7. For further explications of acceptance, see (Bratman 1992; Dellsén 2021; Elgin 2017).
8. I am aware of the fact that the literature on care is wide and complex, and that my analysis glosses over important complexities. My goal in this section is merely to sketch a minimal model of epistemic care that could be deployed to test ChatGPT's performance. For an in-depth analysis of care and caring behavior, see, e.g. (Tronto and Fisher 1990).
9. It has been pointed out to me that my reflections about epistemic care hold only as long as one assumes that the epistemic environment is friendly, or ideal – e.g. it is not polarized, it is not affected by epistemic injustices of various sorts, and so on. It might be that in non-ideal circumstances, epistemically caring behavior looks somewhat different than how I describe it here. I'll leave the exploration of this issue for another project.
10. It could be objected as this point: why should epistemic care involve a sensitivity and responsiveness to another agent's epistemic needs? Suppose I realize that my interlocutor is believing many falsehoods, and I decide to brainwash them and to substitute these beliefs by true beliefs. Wouldn't this be an instance of epistemic care? My answer is: it depends. It depends, among other things, because truth is not the only thing that we value in the epistemic domain. Maybe the falsehoods that my interlocutor is engaging with are not epistemically worthless. Maybe they are 'felicitous' in Elgin's sense (Elgin 2017). Maybe they have a role to play in the tapestry of commitments, beliefs and acceptances that put my interlocutor in the position to make sense of things – at least to some extent. Suppose my 5-year-old child believes that the ball is falling to the ground because of the force of gravity. Correcting her belief and bringing her to believe that objects fall to the ground because of the curvature of spacetime would not be an instance of epistemically caring behavior. Substituting a felicitous falsehood by an unusable, and hence cognitively inert truth would not be in her epistemic interest. The crucial concept here is 'epistemic wellbeing'. Epistemic care, as I suggest conceiving of it, promotes another agent's epistemic wellbeing. Thanks to Thomas Grundmann for raising the objection, and for promoting these reflections.
11. A full account of care should probably enable one to distinguish between cases of paternalism that are instances of care from cases of paternalism that are not. While this is an extremely important project worthy of being pursued, I cannot aspire to solve the issue in this paper. What matters for my purposes is the fact that paternalism and care sometimes overlap.
12. Thanks to Rico Hauswald for pressing me on this point.
13. Epistemic acceptability refers to what is currently accepted by the community of clearly recognized experts.
14. As Floridi and Chiriatti (2020, 683) put it: 'The game of questions (Turing's "imitation game") is a test only in a negative (that is, necessary but insufficient) sense, because not passing it disqualifies an AI from being "intelligent", but passing it does not qualify an AI as "intelligent". In the same way, Ambrogio [the lawnmower] mowing the lawn – and producing an outcome that is indistinguishable from anything [the human] Alice could achieve – does not make Ambrogio like Alice in any sense, either bodily, cognitively, or behaviourally' (emphasis added).
15. But what if we were in the position to 'incorporate' epistemic values into the system? Wouldn't we have authentic epistemic care then? I must admit that I do not really have a clear conception of what this 'incorporation' of values into the system would require. It seems to me that the best one could achieve is an AI system fed only with information of outstanding quality. But even in such a case, the software's seemingly caring response would be 'externally' and not 'internally' motivated.

Acknowledgments

A first draft of this paper was presented at the conference *New Waves in the Philosophy of Epistemic Authority and Expert Testimony* at the TU Dresden in October 2023. I am extremely grateful for the feedback I received in that occasion. I would also like to thank Thomas Grundmann, Rico Hauswald and Pedro Schmechtig for the detailed, enlightening and challenging comments they provided on the manuscript.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Tiroler Wissenschaftsförderung Grant number F.32999/7-2021.

Notes on contributor

Federica Isabella Malfatti is Assistant Professor at the Department for Christian Philosophy of the University of Innsbruck. She works at the intersection between social epistemology and the philosophy of science.

ORCID

Federica Isabella Malfatti  <http://orcid.org/0000-0002-3079-2876>

References

- Audi, R. 2006. "Testimony, Credulity, and Veracity." In *The Epistemology of Testimony*, edited by J. Lackey and E. Sosa, 25–49. New York: Oxford University Press.
- Bourget, D. 2025. "A Phenomenal Theory of Grasping and Understanding." In *Understanding and Conscious Experience: Philosophical and Scientific Perspectives*, edited by A. I. Marasoiu and M. Dumitru. New York: Routledge. <https://philarchive.org/rec/BOUAPT>.
- Bratman, M. E. 1992. "Practical Reasoning and Acceptance in a Context." *Mind* 101 (401): 1–16. <https://doi.org/10.1093/mind/101.401.1>.
- Broncano-Berrocal, F. 2020. "Epistemic Care and Epistemic Paternalism." *Epistemic Paternalism: Conceptions, Justifications and Implications*, 169–182. Lanham, Md: Rowman & Littlefield.
- Cohen, L. J. 1989. "Belief and Acceptance." *Mind* 98 (391): 367–389. <https://doi.org/10.1093/mind/XCVIII.391.367>.
- Croce, M. 2018. "Expert-Oriented Abilities Vs. Novice-Oriented Abilities: An Alternative Account of Epistemic Authority." *Episteme* 15 (4): 476–498. <https://doi.org/10.1017/epi.2017.16>.
- Croce, M. 2019. "For a Service Conception of Epistemic Authority: A Collective Approach." *Social Epistemology* 33 (2): 172–182. <https://doi.org/10.1080/02691728.2019.1577511>.
- Croce, M., and D. Pritchard. 2022. "Education as the Social Cultivation of Intellectual Virtue." In *Social Virtue Epistemology*, edited by M. Alfano, C. Klein, and J. de Ridder, 583–601. New York: Routledge.
- De Regt, H. W. 2017. *Understanding Scientific Understanding*. New York: OUP USA.
- Dellsén, F. 2017. "Understanding Without Justification or Belief." *Ratio* 30 (3): 239–254. <https://doi.org/10.1111/rati.12134>.
- Dellsén, F. 2020. "Beyond Explanation: Understanding as Dependency Modelling." *British Journal for the Philosophy of Science* 71 (4): 1261–1286. <https://doi.org/10.1093/bjps/axy058>.
- Dellsén, F. 2021. *Rational Understanding: Toward a Probabilistic Epistemology of Acceptability* [Preprint]. <http://philsci-archive.pitt.edu/16183/>.
- Diller, A., B. Houston, K. P. Morgan, M. Ayim, and A. Diller. 2018. "The Ethics of Care and Education: A New Paradigm, Its Critics, and Its Educational Significance." 89–104. <https://doi.org/10.4324/9780429496530-10>.
- Elgin, C. Z. 1999. "Education and the Advancement of Understanding." *The Proceedings of the Twentieth World Congress of Philosophy* 3:131–140.
- Elgin, C. Z. 2013. "Understanding's Tethers." In *Epistemology: Contexts, Values, Disagreement*, 131–146. Berlin/Boston: De Gruyter. <https://doi.org/10.1515/9783110329018.131>.
- Elgin, C. Z. 2017. *True Enough*. Cambridge: MIT Press.
- Elgin, C. Z. 2021. "Beyond the Information Given: Teaching, Testimony, and the Advancement of Understanding." *Philosophical Topics* 49 (2): 17–34. <https://doi.org/10.5840/philtopics202149213>.
- Elgin, C. Z. 2022. "Understanding as an Educational Objective." In *Handbook of Philosophy of Education*, edited by R. R. Curren. New York: Routledge.
- Floridi, L. 2019. "What the Near Future of Artificial Intelligence Could Be." *Philosophy & Technology* 32 (1): 1–15. <https://doi.org/10.1007/s13347-019-00345-y>.
- Floridi, L., and M. Chiriatti. 2020. "GPT-3: Its Nature, Scope, Limits, and Consequences." *Minds and Machines* 30 (4): 681–694. <https://doi.org/10.1007/s11023-020-09548-1>.
- Grimm, S. R. 2011. "Understanding." In *The Routledge Companion to Epistemology*, edited by D. P. S. Berneker. New York: Routledge.
- Grimm, S. R. 2014. "Understanding as Knowledge of Causes." In *Virtue Epistemology Naturalized: Bridges Between Virtue Epistemology and Philosophy of Science*, edited by A. Fairweather, 329–345. Cham: Synthese Library, Springer. https://doi.org/10.1007/978-3-319-04672-3_19.
- Grote, T. 2021. "Trustworthy Medical AI Systems Need to Know When They don't Know." *Journal of Medical Ethics* 47 (5): 337–338. <https://doi.org/10.1136/medethics-2021-107463>.
- Hills, A. 2016. "Understanding Why." *Noûs* 50 (4): 661–688. <https://doi.org/10.1111/nous.12092>.
- Jäger, C., and F. I. Malfatti. 2021. "The Social Fabric of Understanding: Equilibrium, Authority, and Epistemic Empathy." *Synthese* 199 (1): 1185–1205. <https://doi.org/10.1007/s11229-020-02776-z>.

- Johnson, C. R. 2022. *Epistemic Care: Vulnerability, Inquiry, and Social Epistemology*. New York: Routledge/Taylor & Francis Group.
- Kelp, C. 2015. "Understanding Phenomena." *Synthese* 192 (12): 3799–3816. <https://doi.org/10.1007/s11229-014-0616-x>.
- Khalifa, K. 2012. "Inaugurating Understanding or Repackaging Explanation?" *Philosophy of Science* 79 (1): 15–37. <https://doi.org/10.1086/663235>.
- Khalifa, K. 2017. *Understanding, Explanation, and Scientific Knowledge*. Cambridge: Cambridge University Press.
- Kvanvig, J. L. 2003. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511498909>.
- Lackey, J. 1999. "Testimonial Knowledge and Transmission." *The Philosophical Quarterly* (1950-) 49 (197): 471–490. <https://doi.org/10.1111/1467-9213.00154>.
- Lackey, J. 2008. *Learning from Words: Testimony as a Source of Knowledge*. Vol. 88. Oxford: Oxford University Press.
- Lipton, P. 2009. "Understanding without Explanation." In *Scientific Understanding: Philosophical Perspectives*, edited by H. W. de Regt, S. Leonelli, and K. Eigner, 43–63. Pittsburgh: University of Pittsburgh Press.
- Malfatti, F. I. 2019. "Can Testimony Generate Understanding?" *Social Epistemology* 33 (6): 477–490. <https://doi.org/10.1080/02691728.2019.1628319>.
- Newman, M. 2012. "An Inferential Model of Scientific Understanding." *International Studies in the Philosophy of Science* 26 (1): 1–26. <https://doi.org/10.1080/02698595.2012.653118>.
- Pritchard, D. 2005. *Epistemic Luck*, 191–220. Vol. 29, Issue 223. Oxford: Oxford University Press UK.
- Pritchard, D. 2009. "Knowledge, Understanding and Epistemic Value." *Royal Institute of Philosophy Supplement* 64:19–43. <https://doi.org/10.1017/S1358246109000046>.
- Pritchard, D. 2014. *Knowledge and Understanding*, 315–327. https://doi.org/10.1007/978-3-319-04672-3_18.
- Pritchard, D. 2020. "Educating for Intellectual Humility and Conviction." *Journal of Philosophy of Education* 54 (2): 398–409. <https://doi.org/10.1111/1467-9752.12422>.
- Seton Bennett, P. 2023. "Care Ethics, Needs-Recognition, and Teaching Encounters." *Journal of Philosophy of Education* 57 (3): 626–642. <https://doi.org/10.1093/jopedu/qhad040>.
- Strevens, M. 2008. *Depth: An Account of Scientific Explanation*. Cambridge: Harvard University Press.
- Strevens, M. 2013. "No Understanding without Explanation." *Studies in History & Philosophy of Science Part A* 44 (3): 510–515. <https://doi.org/10.1016/j.shpsa.2012.12.005>.
- Stueber, K. 2006. *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*. Cambridge: MIT Press.
- Stueber, K. 2019. "Empathy." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2019 ed. <https://plato.stanford.edu/archives/fall2019/entries/empathy/>.
- Tronto, J. C., and B. Fisher. 1990. "Toward a Feminist Theory of Caring." In *Circles of Care*, edited by E. Abel and M. Nelson, 36–54. New York: SUNY Press.
- Watson, L. 2016. "The Epistemology of Education." *Philosophy Compass* 11 (3): 146–159. <https://doi.org/10.1111/phc3.12316>.
- Wilkenfeld, D. A. 2013. "Understanding as Representation Manipulability." *Synthese* 190 (6): 997–1016. <https://doi.org/10.1007/s11229-011-0055-x>.
- Zagzebski, L. 2001. "Recovering Understanding." In *Knowledge, Truth, and Duty*. Oxford: Oxford University Press. <https://doi.org/10.1093/0195128923.003.0015>.