Φιλοσοφια
Volume 27, 2: 2026

EPISTEMIC AIMS AND AI IN EDUCATION PARADIGMS

Robert James M. Boyles*, Mark Anthony Dacela, Joyce Estelle Fungo, Lenor Coreen Aguinaldo
De La Salle University, Philippines

This paper provides added foundational support to the three proffered paradigms of artificial intelligence in education. Specifically, it gives a preliminary analysis of the key epistemic aims present in each of the said frameworks. In order to account for the different ways in which artificial intelligence-related educational strategies have been employed to address issues in learning and instruction, three paradigms have been previously put forward: AI-directed (learner-as-recipient), AI-supported (learner-as-collaborator), and AI-empowered (learner-as-leader). In this paper, it is argued that the different standard views on what ought to be the objectives of education (i.e., philosophical theories that highlight knowledge and truth, critical thinking, intellectual virtues, and understanding) manifest themselves distinctly in these three paradigms. Moreover, apart from epistemically grounding the differing artificial intelligence in education frameworks, this study also demonstrates that more complex models, in principle, could espouse a greater number of epistemic aims.

Keywords: Philosophy of education, epistemology of education, epistemic aims, computer science, artificial intelligence, artificial intelligence in education

INTRODUCTION

In light of the significance of artificial intelligence (AI) in education, several studies have pushed the discussion further, looking into the specific ways the former may be employed to affect the latter (Roll and Wylie 2016; McArthur et al. 2005). For one, Ouyang and Jiao (2021) purport three models of artificial intelligence in education (AIED) to account for the different ways in which AI techniques have been employed to address issues in learning and instruction. These frameworks are as follows: (1) AI-directed, (2) AI-supported, and (3) AI-empowered.

Though Ouyang and Jiao (2021, 2-4) provide the corresponding theoretical underpinnings of the three mentioned paradigms, no further analysis has been given

with regard to the epistemic aims of each. Roll and Wylie (2016, 595), in studying the various AIED research strands, have already raised the question on which learning goal should really be addressed.[1] Furthermore, this also relates with one of the more fundamental questions in education—that, in general, the specific ends of the educational process must be properly identified (Kohlberg and Mayer 1972, 449).

This paper provides added foundational support to the three AIED paradigms mentioned above. Specifically, it gives a preliminary analysis of the key epistemic aims present in each of these frameworks. To demonstrate the said goals, the subsequent section first summarizes Ouyang and Jiao's (2021) view regarding the different AIED models. The following section, on the other hand, discusses the standard philosophical theories on the epistemic aims of education, particularly those that underline (1) knowledge and truth, (2) critical thinking, (3) intellectual virtues, and (4) understanding. The next section, then, identifies the apparent epistemic goals in the three AIED models. Apart from philosophically grounding the said frameworks, this section also demonstrates that more complex paradigms, in principle, could espouse a greater number of epistemic aims. The last section ends with some concluding remarks, as well as a few recommendations for future research.

THREE AI IN EDUCATION PARADIGMS

Paradigm One: AI-directed or Learner-as-recipient

The first AIED framework, AI-directed, is characterized by Ouyang and Jiao (2021, 2-3) as following a learner-as-recipient model. This is because, in this paradigm, the "AI represents the domain knowledge and directs the learning processes while the learner acts as [a recipient] of AI service to follow the specific learning pathways" (Ouyang and Jiao 2021, 2). So, this AIED paradigm considers learners as recipients of knowledge; they are, for instance, logically and incrementally introduced to concepts, provided feedback instantaneously, and are positively reinforced by AIs with programmed instructions. Ouyang and Jiao (2021, 2) further explain that a standard example of programs that realize the AI-directed model are Intelligent Tutoring Systems.

As per Ouyang and Jiao (2021, 2), among the standard implementations of Intelligent Tutoring Systems that follow the first AIED model include the ACT Programming Tutor and Stat Lady. In a nutshell, the ACT Programming Tutor is a problem-solving system that enables students to learn and develop short programs in Lisp, Pascal, or Prolog (Corbett 2000, 31). As regards the modules for these programming languages, Corbett and Bhatnagar (1997, 244) explain that they employ a language-specific cognitive model that "enables the tutor to trace the student's solution path through a complex problem solving space, providing feedback on problem solving actions and, if requested, advice on steps that achieve problem solving goals."

As regards Stat Lady, Ouyang and Jiao (2021, 2) hold that this program teaches statistics to students by systematically introducing curriculum content in a fixed order, requiring them to first solve a set of predefined problems. Once the said tasks are

accomplished, a student is assumed to have mastery of them and is allowed to move forward to the subsequent step.

In relation to its theoretical foundation, Ouyang and Jiao (2021) explain that AI systems under the AI-directed framework make use of behaviorist strategies and inherit the characteristics of Skinner's (1958) teaching machines. So, it may be said that AIs that subscribe to the AI-directed paradigm focus more on rote learning, since behaviorist techniques have been generally "regarded as facilitating surface learning rather than deep understanding" (Buchanan 1998, 75).

Paradigm Two: AI-supported or Learner-as-collaborator

Ouyang and Jiao (2021, 3) state that the second AIED model, AI-supported, typically embodies the learner-as-collaborator concept. This entails that AIs under this paradigm act more as supporting tools in the learning process (i.e., rather than having full control of it), and learners become collaborators with these technologies. As compared to the first paradigm, the AI-supported framework also facilitates learner-centered learning better, since it fosters the interaction and collaboration between AIs and their learners.

As explained by Ouyang and Jiao (2021, 3), the typical implementations of AIs that follow the AI-supported framework, like dialogue-based tutoring systems or exploratory learning environments, foster system-learner interactions. This means that these kinds of AI systems gather and analyze data from learners in order to arrive at an accurate picture or understanding of their learning status. At the same time, however, learners could also explore and communicate with these AIs in order to have a better understanding of their decision-making process, which further aids students in prospectively coming up with better choices for learning.

In highlighting the capacity of learner-as-collaborator-based AIs to collect and analyze data from learners, Ouyang and Jiao (2021, 3) also cite examples of systems that exemplify this ability, like those proffered by Stamper (2006) and Käser et al. (2017). For one, Stamper's (2006, 206) proposed system has the capability of continued refinement whenever new data is introduced. Käser et al. (2017, 450), on the other hand, make use of dynamic Bayesian network models to improve the accuracy of representing the knowledge of students, specifically by considering the prerequisite skill hierarchies of learners.

Furthermore, with regard to the capability of AI-supported systems to enable learners to explore and communicate with AI programs (i.e., to better make sense of their decision-making process), Ouyang and Jiao (2021, 3) identify the exploratory learning environment QUE, short for "QUerying the Expert" (Metzler and Martincic 1998), as an example. This program was created so that students could explore and check the discrepancies between their incorrect answers in an Intelligent Tutoring System and what an AI considers as the correct line of reasoning. So, even if AI-supported AIs may still be considered as Intelligent Tutoring Systems, they are quite different from systems under the AI-directed paradigm, since there is continuous refinement of learning as students interact with the AI programs, allowing the former to better explore, grasp, and understand why the latter gave specific suggestions.

As regards the theoretical underpinnings of the AI-supported framework, Ouyang and Jiao (2021, 3) explain that it heavily relies on the cognitive and social constructivist views of learning. They further clarify that, under the said views, the assumption is that learning transpires in socially-situated contexts, specifically when learners come in contact and interact with information, technology, and other people. Learner-centered education may also be traced back to constructivist developmental theories (Mascolo 2009, 4). Other studies, like that of Powell and Kalina (2009), have already underscored the value of employing both cognitive and social constructivism in the context of, say, classroom learning.

Paradigm Three: AI-empowered or Learner-as-leader

The third AIED model, AI-empowered or learner-as-leader, is depicted by Ouyang and Jiao (2021, 3) as heavily focusing on learner agency. This is because the said framework considers AI systems as tools that could enhance and augment human intelligence. Ouyang and Jiao (2021, 4) explain that advanced AI techniques, like brain-computer interfaces, machine learning, and deep learning, bring about sustained data collection and analysis, which naturally yields more accurate, transparent, and interactive data. To better understand the main focus of AI-empowered systems, Ouyang and Jiao (2021, 3-4) also cite a number of examples related to the third AIED paradigm.

As highlighted by Ouyang and Jiao (2021, 3), synergetic collaborations in complex systems are fostered by AIs that subscribe to the learner-as-leader paradigm, like those purported by Hoc (2000) and Riedl (2019). Following the idea that all forms of relationships between humans and machines may be deemed as instantiations of human-computer interactions, Hoc (2000, 833-834) defends the view that a framework of cooperation, also known as "human-machine cooperation" (HCR), should be adopted in dynamic situations.

Meanwhile, guided by the idea that intelligent machines with the capacity to come up with some autonomous decisions (i.e., based on a given set of goals) may be treated as agents, Riedl (2019) advocates for the development of human-centered AIs. The latter espouses the idea that intelligent systems should be developed in light of the broader picture that they would eventually be part of a system composed of human stakeholders (e.g., users, clients, etc.).

For Ouyang and Jiao (2021, 3-4), the ideas offered by Hoc (2002) and Riedl (2019), along with proposals related to human AI-collaborations (Hwang et al. 2020) and human-centered AIED (Yang et al. 2021), are AI approaches that also consider human conditions, expectations, and contexts. Hwang et al. (2020, 2), for example, argue that AI systems in education are not only limited to being deemed as intelligent tutors, tutees, or partners, but also potential policy-making advisors for, say, policy building in education.

Yang et al. (2021), on the other hand, highlight the importance of designing human-centered AIs, since they could affect human welfare and impact society. For instance, as regards the field of education and smart learning environments, Yang et al. (2021, 3) forewarn that AIs "must not only focus on performance; human feelings and outcomes should be the main concern." So, it could be said that systems that would

fall under the AI-empowered paradigm may be deemed as the target or aim of those using AI in education.

In terms of the theoretical foundations of the AI-empowered paradigm, Ouyang and Jiao (2021, 4) argue that it is primarily guided by complexity theory. In general, this theory draws from different fields of inquiry, especially key concepts from chaos theory (Mason 2008, 5). Morrison (2008, 19) further explains that complexity theory is characterized by adaptation, development, evolution, and change. In light of the notion that education is a complex adaptive system, Ouyang and Jiao (2021, 3) claim that there should be a synergetic collaboration between multiple learners, instructors, and technologies, including information, in this kind of system.

EPISTEMIC AIMS OF EDUCATION

Outside the philosophy of education, the subject of epistemic aims has also been a widely debated topic (Robertson 2010; Watson 2016), wherein the question is framed in a slightly different manner. Given that epistemology generally refers to the theory of knowledge,[2] the epistemic domain may be considered the realm of knowledge-related things, broadly speaking. *Justification*, for instance, is an epistemic subject precisely because knowledge is often regarded as a true belief that is strongly justified.[3] *Truth* comes into play because, for many theories of knowledge, a subject knowing any truth-evaluable proposition would require that the statement in question be true, rather than false (i.e., one cannot really know a proposition if it is incorrect to begin with). The virtue-theoretic approach to knowledge, meanwhile, would maneuver the analysis to a more Aristotelian point of view in that, in order to become a good knower, a subject must habitually cultivate intellectually virtuous traits.[4] It is no surprise then that education, being an enterprise that deals with a subject undergoing the learning process, is, by extension, an epistemic concern.

The connection between epistemic aims and education, thus, becomes more apparent given the fact that the latter goes hand in hand with knowledge. In the aforementioned AIs and their interaction with human subjects, one could wonder about what it is in the whole process that renders an education complete, like if it were approached from an end-driven point of view, if not perhaps successful. In other words, the primary consideration is when one can say that a student has actually learned what they must learn. By implication, it may be asked whether there really is an ideal good (or goods) that a learning subject ought to aim at. The negative alternative to this, of course, would state that the learning process does not really have an aim to start with. Note that this study will not address the latter problem. For the purposes of the present paper, it is assumed that education has an epistemic aim.

Whether one aims for knowledge when undergoing the process of learning is the question that ties the two endeavors (i.e., education and epistemology) together. Such aims could be deemed epistemic because, outside of knowledge and knowledge-related things, there are many other good reasons why a learner may want to be educated. A learner could aim to be educated in a formal setting for pragmatic reasons, like the goal to eventually become employable. On the other hand, one could, for instance, wish to be educated to play the piano in order to impress a lover, in which

case the reason would be prudential or self-interest. Addressing the question of epistemic aims requires something much more different than those of prudential or pragmatic aims, since the said endeavor seeks to address what it is among epistemic goods, beyond individual interests, desires, and needs, that must be cultivated to become better cognitive subjects. One important social institution that allows learners to achieve this is education. In relation to the paradigms in the beginning sections, the normative authority that epistemic goals have on individuals ought to provide the theoretical grounding for what makes any AIED framework worthy of attention for educators and learners alike.

The subsections that follow concisely expose the existing contenders for the epistemic aims of education, serving as signposts or practical summaries to these competing views. As said, the question regards what epistemic goods the learning process is aimed at. To address this, one may turn to the available positions in the epistemology of education, with the hope of carrying them over to discussions about AI paradigms. The next subsections would cover the following contenders: (1) knowledge and truth, (2) critical thinking, (3) intellectual virtues, and (4) understanding.

Knowledge and Truth

To understand Goldman's (1999) motivations in proffering a theory of truth, it would help to first identify the area in epistemology from which his view on the aims of education may be drawn. In *Knowledge in a Social World*, Goldman (1999) maintains that collective epistemic agents, as well as social institutions, ought to actively take part in the proliferation of truth—a view that has been associated with social epistemology (SE).

It may be said that Goldman (1999) has been guided by the SE framework in developing his position on the aims of education, since this also extends to the goals of other social institutions, like science and law. Also, SE espouses a catalogue of focal themes that are considerably useful in the realm of education (e.g., treating the problem of testimony as an epistemic issue, underscoring the issue of the expertise of educators, problematizing what epistemic aims education should attain, etc.).

For Goldman (1999), the principal goal of education is to cultivate knowledge production among learners. Notice that what makes such a veritistic (or truth-oriented) view is how Goldman (1999) first accounts for knowledge. Like many post-Gettier epistemologists who have investigated the nature of knowledge, Goldman (1999) holds the view that knowledge is not equivalent to justified true belief. In contrast, he contends that knowledge could be defined as true belief (i.e., at its most basic level).

Goldman (1999) defends a reliabilist conception of knowledge, which subscribes to the idea that true beliefs are basically acquired through a reliable cognitive belief-forming process. So, this view differs from other accounts of knowledge, since it considers the process by which the belief was arrived at. The end result of the said process is a true belief, which makes the said view veritistic.

Critical Thinking

For philosophers, like Marabini and Moretti (2020) and Siegel (2005), critical thinking should be the epistemic aim of education. For one, Siegel explains that "education should strive to foster, not (just) true belief, but (also) the skills, abilities and dispositions constitutive of critical thinking, and the rational belief generated and sustained by it" (2005, 347). In order to defend the said view, Siegel further argues that Goldman's (1999) position is flawed by citing two problems.

First, Siegel (2005, 346) holds that Goldman (1999) merely treats critical thinking as something instrumentally valuable to the pursuit of truth. Second, he further contends that Goldman (1999) promotes the roles of testimony and trust in education, as well as in epistemology, to demerit critical thinking. It must also be highlighted here that Siegel (2005) does not consider the promotion of truth as problematic on its own. Instead, he deems critical thinking to be substantial for the pursuit of truth.

In relation to the first problem, Siegel (2005) raises the plausibility of lucky guesses, which are beliefs that are not rationally held, yet are otherwise true. He then clarifies that critical thinking is not merely instrumental to truth, but on equal footing with it in terms of being an educational goal. The second problem, on the other hand, is motivated by the fact that Goldman (1999) holds the view that the cultivation of critical thinking could be counterproductive, since education heavily relies on trust and testimonial accounts of educators. Siegel, in contrast, refutes this, arguing that "young children typically do not have testimony-independent reasons for trusting their teachers" (2005, 362).

Furthermore, Siegel claims that, "[f]or older children, and adult students, there is virtually always such independent reason for trust" (2005, 362). Here it may be asserted that, in many cases, the said reasons come from critical thinking, which also highlights the notion that education should result in both true and rational belief—a term Siegel (2005) uses interchangeably with critical thinking.[5]

Intellectual Virtues

For Baehr (2013), education must aim at fostering intellectual virtues. Borrowing Baehr's (2013, 229) own words, the following argument could, thus, be formulated and reconstructed to further expose his stance:

> [i] [E]ducation should aim at fostering 'lifelong learning'.
> [ii] [W]e can think of intellectual virtues as the *personal* qualities or characteristics of a lifelong learner.
> [iii] Being a lifelong learner also requires being *curious* and *inquisitive*. It demands *attentiveness* and *reflectiveness*… it also requires intellectual *determination*, *perseverance*, and *courage*.
> [iv] In other words, being a lifelong learner is largely constituted by the possession of various intellectual virtues.
> (Baehr 2013, 249)

Proposition (i) may be deemed as a generally accepted principle—rarely do we seek the fruits of education to last us only in the short term. Proposition (ii) characterizes the relationship between intellectual virtue and lifelong learning. This relationship centers on character embodiment, wherein the presumed, ideal lifelong learner realizes a specific set of virtues of the mind. Notice here that such a depiction resembles the Aristotelian view of moral virtues as qualities that make up the character of an ideal moral agent (Baehr 2013, 258).

So, while moral character virtues address the issue of what it means to live a (morally) good life, intellectual character virtues answer the question of what it means to live an intellectually prosperous life. Thus, what follows is the question on *which* sorts of intellectual virtues must be cultivated?

Proposition (iii) says that curiosity, inquisitiveness, attentiveness, reflectiveness, intellectual determination, intellectual perseverance, and intellectual courage should be able to address the said query. In defense of these virtues, Baehr (2013, 249) claims that to be a lifelong learner, one should be strongly committed to learning, and this strong commitment is demonstrated by the specified virtues. Thus, combining the aforementioned virtues inevitably results in lifelong learning—a conclusion encapsulated in proposition (iv). Similar to Siegel's (2005) position, Baehr also thereby adopts a pluralistic view—in fact, more so than the former.

Understanding

Several thinkers have also identified understanding as a viable option in terms of being an epistemic aim of education (Pritchard 2013; Fungo and Dacela 2022). In order to better grasp this idea, it would be helpful to first highlight the various types of understanding.

Analytic traditions in epistemology have typically considered three differing types of knowledge: acquaintance knowledge, knowledge-how, and propositional knowledge. A large majority of the literature within mainstream epistemology in recent years—from 1963 onwards[6]—has focused on providing an analysis for the third kind. Although they do not perfectly mirror these types of knowledge, there are also three contrasting notions of understanding (Kvanvig 2003, 188-191), which are as follows: (1) propositional understanding, (2) understanding-why, and (3) objectual understanding. Zeroing in on understanding-why, Carter and Gordon (2014, 3) flesh out Duncan Pritchard's position, explaining that this concept of understanding is (1) factive (2) it needs "good reflectively accessible grounds" as a support, and (3) it is unlike propositional knowledge as it is resilient from epistemic luck.

In light of the assumption that understanding-why is reconcilable with specific forms of luck, like environmental luck, Pritchard argues for certain possibilities, wherein even if an "agent lacks knowledge due to the inhospitable epistemic environment, it is not obvious that he also lacks *understanding*" (Carter and Gordon 2014, 3). Pushing the argument further, Pritchard (2013) thus advances the notion that understanding should be the fundamental epistemic goal of education. He further contends for such by appealing to the notion of epistemic value, asserting that understanding may be arrived at via enhancing an epistemic agent's cognitive abilities.

Additionally, Pritchard (2013, 241) emphasizes that strong cognitive achievements are of final value. This is in contrast to only ascribing instrumental value to such, wherein the value of the epistemic good just lies in its capacity to achieve another goal. For instance, it could be said that critical thinking is only valued due to its capacity to arrive at truth. If one grants this, then it may be said that such does not possess inherent value.

If the sole function of critical thinking is to act as a stepladder towards truth, then it may also be claimed that it is only of instrumental value, since it merely bridges the gap between the epistemic agent and ultimate end goal, which is truth in this particular context. However, since understanding is a strong cognitive achievement, Pritchard (2013) deems it to be of final value, rather than being instrumentally valuable only.

AIED FRAMEWORKS AND THEIR EPISTEMIC AIMS

Epistemic Aims of the AI-directed Paradigm

With regard to Ouyang and Jiao's (2021) first AIED framework, AI-directed, it may be argued that this model relates with Goldman's (1999) idea that truth should be deemed an epistemic aim of education. Such is the case on account of two things. First, the AI-directed paradigm seems to prescribe Goldman's (1999) reliabilist conception of knowledge, which holds that knowledge as true belief is acquired through a reliable cognitive process.

As regards Intelligent Tutoring Systems, for instance, it may be said that one of the primary motivations in developing and employing these systems is that they are reliable enough to yield a high proportion of true beliefs. This is evident based on how the design of both the ACT Programming Tutor (Corbett 2000) and Stat Lady (Shute et al. 1996) also emphasize adaptive learning (i.e., with the goal of perfecting the system to guarantee results)—the ACT Programming Tutor has its system reliant on modules in directing the learning process, while Stat Lady depends on curricula that an AI introduces systematically in a fixed order.

Second, since the AI-directed model theoretically hinges on behaviorism, wherein its techniques consistently facilitate rote learning better, it does not seem directed towards the other aims that require more than mere surface learning. Both the ACT Programming Tutor and Stat Lady programs place much emphasis on the mastery of content, and not so much on virtue cultivation, critical dispositions, or understanding-why. Critical thinking, intellectual virtues, and understanding appear to have structures different from true beliefs, which the AI-directed model seems incapable of facilitating. For one, most of these other epistemic goals arguably necessitate constructive interaction between AIs and their learners.

However, it should be pointed out that claiming that truth is an epistemic aim of the AI-directed paradigm does not necessarily mean that it is the only goal of education. Also, to say the AI-directed paradigm merely aims at truth is not necessarily a bad thing. For Greco (2021, 50), "knowledge is commonly understood as being (partially) constituted by truth."

In light of the different challenges proffered against the standard philosophical account of knowledge (i.e., justified true belief), like Gettier's (1963) own objections, it may be said that almost all attempts to salvage the said definition have been mainly directed towards the aspect of justification; both truth and belief conditions seemingly remain unchallenged in discussions on the constituents of knowledge. So, it appears that truth is something attainable with, say, formal or top-down instruction.

As also discussed earlier, Ouyang and Jiao's (2021, 2-3) AI-directed framework has been characterized as following a learner-as-recipient model. So, in this paradigm, AI systems are considered the domain of knowledge, directing the learning processes, while learners act as the recipients of knowledge. It should be recalled that, as also mentioned above, there is a close relationship between true beliefs and knowledge—one that is of a constitutive kind (i.e., that true beliefs are contained in propositional knowledge according to most accounts) (Greco 2021). Thus, it may be claimed that, whenever learners are logically and incrementally introduced to certain concepts by means of the said framework, they are actually being inculcated with true beliefs.

Furthermore, if one grants that the epistemic goal present in the AI-directed paradigm, which theoretically is grounded on behaviorism, is, in fact, the acquisition of true beliefs, then it may be argued that the said content needs to be guided, if not be composed, of true beliefs, aiding students in aptly translating their learnings afterward.

In order for learners to achieve true beliefs, which closely relate to knowledge in the AI-directed paradigm, it has been said that the following are deemed necessary: (1) learners should be epistemically committed to believing a proposition, and (2) the said proposition must also be true. A number of research on the epistemology of education have also given enough attention to the idea of epistemic authority, as well as trust in teaching and learning. Goldman (1999, 367), for one, already covered these in passing, though much of his focus was on which truths are to be taught in curricula. Zagzebski (2012), on the other hand, also highlights this, but in a relatively more extensive manner.

Zagzebski (2012, 37), for instance, explains that "trust includes an element of dispelling doubt with the awareness that doing so makes one vulnerable," and it is also "an attitude opposite to doubt … that can be appropriately directed towards nonpersons." So, it may be argued that, in order to accomplish the epistemic goal of truth, learners must first develop trust in a system, like a human teacher or an AI program, and consider such an epistemic authority. Also, Zagzebski (2012, 139) clarifies that "belief *can* be coerced in the sense that it is possible to impose sanctions on epistemic subjects who fail to take beliefs on authority." Note that the depiction given above does not necessarily entail that a coercive teaching strategy should be endorsed.

In light of the said considerations, it may be argued that AI systems that subscribe to Ouyang and Jiao's (2021) first AIED paradigm could also be viewed as epistemic authorities by their learners, thereby facilitating the attainment of truth for them. Recall that, for Ouyang and Jiao (2021, 3), AIs that follow the AI-directed model are considered the primary drivers of the entire learning processes, taking on the role as directors, while their learners or students act as recipients.

Considering that typical AI systems that espouse the AI-directed paradigm are arguably not that sophisticated, like Intelligent Tutoring Systems, it may be said that

other epistemic aims apart from truth are not that prominent in such. As pointed out earlier, the epistemic aims of critical thinking, intellectual virtues, and understanding seem to have structures quite different from true beliefs, which AI-directed systems would be incapable of facilitating, as the said epistemic goals necessitate constructive interaction between an AI system and its learners.

Epistemic Aims of the AI-supported Paradigm

In terms of the epistemic goals that could be attributed to Ouyang and Jiao's (2021, 3) second paradigm, AI-supported, recall that AI systems that subscribe to this model are primarily utilized as supporting tools in the learning process. So, it seems contrived to suggest that learners should also treat them as epistemic authorities (i.e., in the same manner as how AI systems are deemed in the AI-directed framework). Nevertheless, it could be said that AI-supported systems are quite capable of instrumentally arriving at truth by first developing critical thinking.

The simple recognition of educators as epistemic authorities (i.e., while building trust in learners) is arguably insufficient in fostering critical thinking. For one, achieving this epistemic aim might even be contradictory in some cases due to the very nature of critical thinking itself. The latter requires learners to at least have the capability to question, if not reject, information provided to them in light of, say, sufficient evidence—a key feature that does not seem to be present in AI-directed systems. In contrast, those that follow the AI-supported paradigm seem to support this given its emphasis on the AI-learner interaction, as well as collaborations that are sustained by dialogue and learning explorations.

In the AI-supported model, the learning relationship between the AI and learner makes the latter an active participant in the learning process; the learner is not a mere passive recipient. This active role carries with it responsibilities of the sort that require the learner to make decisions that would impact the whole learning process. This is only possible because AI-supported systems allow the learner to be critical (i.e., not only of the learning process in which they participate, but of their own participation and learning situation).

The learning environment espoused by programs that subscribe to Ouyang and Jiao's (2021, 3) second AIED model only work if the situation, knowledge, and skills of the learner are well represented in the system such as in the examples of Stamper (2006) and Käser et al. (2017), and when learners have the freedom to gather relevant information to examine both the content and implementation of the learning program. Such things enable students to ask "what if" and "why not" type of questions, which allows them to think about and make critical decisions that, in due course, would shape their learning experience. This is also manifested in Metzler and Martincic's (1998) QUE, since it empowers its learners to explore and check the discrepancies between their incorrect answers and what the AI system considers as the correct line of reasoning.

So, in AI-supported systems, there is continuous refinement of learning as students interact with the programs, which gives learners a better chance of understanding why they were provided particular suggestions by the latter. The

learner, thus, becomes a more active participant in the learning process, since AI-supported systems also allow and invite learners to be critical of the learning process, as well as their own participation and learning situation.

Recall that Ouyang and Jiao (2021, 3) explain that the AI-supported model typically embodies the learner-as-collaborator concept. So, AI systems that follow this framework stand more as supporting tools in the learning process, rather than having full control of it; learners could collaborate with these kinds of technologies and are not directed by them. Since the AI-supported paradigm facilitates learner-centered education better, collaborative interactions between AI systems and their learners also take place, which may pave the way for critical thinking.

Siegel (2010, 143) maintains that, in order to foster critical thinking, educators should strive to enable and encourage learners to "think for themselves, competently and well, rather than to deny them the fundamental ability to determine for themselves, to the greatest extent possible, the contours of their own minds." It may, thus, be said that critical thinking guides and invites learners to go beyond the passive reception of truths, allowing for the possibility of questioning conventional epistemic authorities. Rather than simply accepting truths by means of, say, direct or top-down instructions, a critical thinker arrives at them independently. So, in light of its collaborative nature, truth is arguably achievable in the AI-supported paradigm as well.

Additionally, as mentioned earlier, the theoretical underpinnings of the AI-supported framework heavily rely on the cognitive and social constructivist theories of learning (Ouyang and Jiao 2021, 3). In view of this, the AI-supported model, thus, assumes that learning transpires in socially-situated contexts, wherein learners come in contact and interact with information, technology, and other people. It could, then, be further highlighted that the type of learning environment promoted by the said AIED framework is one that is conducive to critical thinking, given that it is more sensitive and responsive to the unique learning needs of the participants.

Following Hoover (1996), Amineh and Asl (2015, 11) hold that one key implication of constructivism for teaching is that it cannot be deemed as a mere transmission of knowledge, since constructivist teachers act more as guides to their students. They also claim that "[c]onstructivist teaching just promotes learners' motivation and critical thinking, and encourages them to learn independently" (Amineh and Asl 2015, 11). In light of such, it may be said that the AI-supported model arguably promotes critical thinking as this framework also grounds itself with constructivist views.

However, it must also be pointed out that if one adheres to Goldman's (1999) thesis that critical thinking espouses an instrumental relation towards the attainment of truth, it may be said that fostering critical thinking could also lead towards the achievement of truth. In contrast to the AI-directed model, wherein there is a direct transference of truth from AI systems to learners, the AI-supported framework advances critical thinking first, and if practiced habitually, truth may follow afterward as a result of such. So, to say that the AI-supported framework has critical thinking as its goal does not imply that it does not aim for truth at the same time.

Epistemic Aims of the AI-empowered Paradigm

As regards the epistemic aims present in the Ouyang and Jiao's (2021, 3) third AIED paradigm, AI-empowered, recall that AI systems that follow this model highlight the agency of learners, and this makes them better equipped for fostering responsibilist intellectual virtues in contrast to reliabilist ones. The difference between these two may be understood by bearing in mind that, for virtue reliabilists, intellectual virtues are (1) reliable, (2) "hard-wired faculties, or acquired skills," and (3) they "do not require acquired motivations or acquired dispositions of action" (Battaly 2014, 181-182). In contrast, virtue responsibilists (1) "disagree about whether the intellectual virtues require reliably achieving true beliefs," and, for them, intellectual virtues are (2) considered "praiseworthy because they are actively acquired" and (3) "require acquired motivations and acquired dispositions of action" (Battaly 2014, 182-183).

Responsibilist intellectual virtues are modeled after, if not structurally analogous, to Aristotelian moral virtues (Battaly 2014). So, this entails that their acquisition is reliant on person-specific qualities. As discussed earlier, among the said virtues include open-mindedness, attentiveness, curiosity, and intellectual humility, to name a few.

Evidently, each intellectual virtue is different from the other, but what is common among them is that they are all necessarily acquired, requiring dispositions of action and motivation (Battaly 2016, 168). In particular, these virtues necessitate a certain motivation to care about truth, if not about the "cognitive contact with reality" in a broader sense (Zagzebski 1996, 167, as cited in Battaly 2016, 168). It goes without saying, then, that, similar to critical thinking, intellectual virtues identify truth as one of its end-goals.

In terms of the issue of how intellectual virtues are fostered, Battaly (2016) offers three general methods to achieve this, namely: Formal instruction, exemplars, and practice. Also, Battaly recommends that educators...

> …(1) use formal instruction to explain the responsibilist virtues; (2) use exemplars to further elucidate individual responsibilist virtues; (3) provide opportunities to practice identifying virtuous actions, emotions, and motivations; (4) use exemplars to help our students experience virtuous emotions; (5) provide opportunities to practice performing virtuous actions and having virtuous emotions and motivations; and (6) provide opportunities to practice virtuous perception. (Battaly 2016, 174)

In these six steps, the active engagement of learners with learning tools, as well as the exercise of their agency, are quite integral. Out of the three AIED paradigms proposed by Ouyang and Jiao (2021), it could be argued that the AI-empowered model is the one that most embodies these steps. This type of agency is more possible in systems such as those purported by Hoc (2000), wherein there is cooperation between the AIs and their users.

What is crucial in the idea of a dynamic cooperation between humans and machines is a system that also has the capacity to understand the complexity of the sociocultural situation of the learner, such as that espoused by Riedl (2019) and Yang et al. (2021). So, it may be argued that the design and development of human-centered

learning technology could possibly cultivate or deter virtues, which extends its impact to human welfare and society.

Recall that, for Ouyang and Jiao (2021, 3), the AI-empowered framework considers AI systems as tools that could enhance and augment human intelligence. They also further explain that, in order to produce more accurate, transparent, and interactive data, AI techniques (e.g., brain-computer interfaces, machine learning, deep learning, etc.) must focus on sustained data collection and analysis (Ouyang and Jiao 2021, 4). In light of these things, it could be said that intellectual virtues are more realizable in this model (i.e., as compared to the AI-directed and AI-supported paradigms). Consider, for instance, the intellectual virtue of open-mindedness.

One could argue that the likelihood of attaining open-mindedness requires that learners first have access to accurate and reliable data, which AI-empowered systems put more emphasis on than AIs that follow the learner-as-recipient and learner-as-collaborator frameworks. Additionally, notice that certain AI techniques, like brain-computer interfaces, machine learning, and the like, seem to enable the proliferation of open-mindedness better (i.e., in comparison to AI-directed's Intelligent Tutoring Systems and AI-supported's dialogue-based tutoring systems, exploratory learning environments, and the like). Moreover, perhaps more importantly, remember that AI-empowered systems highlight the importance of technologies being more mindful of human conditions, expectations, and contexts, which are key things in terms of the virtue of open-mindedness.

In relation to the theoretical underpinning of the AI-empowered model, which is complexity theory (Ouyang and Jiao 2021, 4), it may be argued that ascertaining intellectual virtues could be the epistemic aim of this framework, which fits such a characterization. Battaly's (2016, 174) idea that there should also be opportunities to "practice virtuous perception," "practice performing virtuous actions and having virtuous emotions and motivations," and so on are better realized in such complex environments that facilitate synergetic collaborations, which, for Ouyang and Jiao (2021, 3), must involve multiple learners, instructors, technologies, and information.

In light of the idea that education is a complex adaptive system, learners should, thus, be better equipped to traverse such kinds of dynamic environments. It may be said that Battaly's (2016, 168) view that intellectual virtues must be motivationally grounded by "cognitive contact with reality" also answers the same call. Actually, it may be further presumed that, if intellectual virtues are imparted and habitually practiced by learners, truth and understanding, if not critical thinking as well, may also be achieved as end results, which better enable learners to adapt to certain complexities.

There seems to be no certain way to directly achieve understanding, since its indicators are only internally measurable within the cognitive life of learners. In light of its relatively more complex structure, especially as compared to truth and knowledge, it could be maintained that trying to account for understanding would necessitate one to first foster epistemic goods that would later be instrumental to its achievement. For example, Kvanvig (2003, 149) explains that the intellectual virtue of curiosity is instrumental for one to arrive at understanding, which further entails that the latter may be achieved by first practicing and realizing the former.

In view of the foregoing discussions, what could be drawn from this present analysis is that the more complex the AIED framework evolves, the number of epistemic goods it makes accessible to the learner also increases. The relatively cruder paradigm, AI-directed, seems to only foster less epistemic goods, while the advanced AI-empowered model opens the possibility of more goods to be realized. So, as AIs advance and evolve, the said epistemic aims become more pronounced as well.

CONCLUSION

As demonstrated above, the three AIED models purported by Ouyang and Jiao (2021) espouse disparate epistemic aims. For one, it seems that it is possible that various epistemological goals are present in a single framework at the same time. Also, it may be said that more complex AIED paradigms, in principle, could support a greater number of epistemic aims.

More importantly, as Ouyang and Jiao (2021) have previously hinted, one may surmise that what goes with each advancement in technological complexity, perhaps including epistemic achievement as demonstrated in the present study, is the social stake—human welfare and development. After all, epistemic goods are social goods; a society flourishes when its members are truth seekers, critical thinkers, and intellectually virtuous. This reinforces Ouyang and Jiao's claim that AIED is more than just about the implementation of AI technology, since it integrates "pedagogical, social, cultural, and economic dimensions" (2021, 5). In light of this, hopefully, this account could provide further assistance to AI researchers and developers.

By knowing that specific epistemic aims become more explicit in much advanced systems, perhaps this could better guide AI developers in modeling teaching and instructional programs in the future. Additionally, considering the findings of this study, hopefully, researchers in the fields of AI, education, history, and the philosophy of education, among others, may be further guided in tracing, fleshing out, and accounting for how the different AI systems have been employed in addressing issues related to learning and instruction.

Prospectively, looking into how the three AIED paradigms, along with their correlative epistemic aims, relate to other facets of learning and instruction would be quite interesting. Consider, for instance, the aspect relating to teacher learning, which could be defined as "the process by which teachers move towards expertise" (Kelly 2006, 514). For one, Cochran-Smith and Lytle (1999) offer three conceptions of teacher learning, namely: knowledge-*for*-practice, knowledge-*in*-practice, and knowledge-*of*-practice.

According to Cochran-Smith and Lytle (1999, 250), the knowledge-*for*-practice paradigm refers to the formal knowledge and theories used to improve one's practice. On the other hand, the second model, knowledge-*in*-practice or practical knowledge, pertains to knowledge gained from experience and practice. Lastly, knowledge-*of*-practice highlights the importance of the application of one's learning, which eventually leads to further knowledge. The knowledge-*of*-practice framework involves learning through collaboration, systematic analysis, and interpretation of knowledge between teachers and students. So, it is quite intriguing how these teacher learning

paradigms further relate to Ouyang and Jiao's (2021) AIED models, especially since the epistemic aims of the latter have already been exposed.

Furthermore, it should be pointed out that the AI-related technologies cited by Ouyang and Jiao (2021) may still be classified to fall under weak AI—a view that generally refers to AIs that only demonstrate human-like behaviors (Bringsjord and Govindarajulu 2018). This conception, however, contrasts with the strong AI view, which subscribes to the belief that computers may also be used to develop artificial minds that possess consciousness comparable to that of human beings.

Among the projected advanced AIs include machines that possess Artificial General Intelligence (Goertzel 2014), which are systems with broad generalization capacities. There are also other AI-related proposals that are relatively more extreme than the typical human-computer interactions, like the gradual replacement of brain neurons with, say, silicon chips that have the same functional characteristics (Chalmers 2010). So, the concept of strong AIs might still be too abstract at present, but one could not help but be curious about how these technologies might possibly affect the mentioned AIED frameworks, including their epistemic aims, if not also education, in general.

NOTES

1. Other scholars have also proffered their own way of accounting for the trends in AIED. For example, see Roll, Ido and Ruth Wylie. 2016. "Evolution and Revolution in Artificial Intelligence in Education." *International Journal of Artificial Intelligence in Education* 26: 582-599.

2. Epistemology has often been called the "theory of knowledge." See Melchior, Guido. 2023. "The Value of Knowledge and Other Epistemic Standings: A Case for Epistemic Pluralism." *Philosophia* 51: 1829-1847.

3. Christoph Kelp explains that there are exceptions to this view such as *knowledge-firsters*, who "typically hold that knowledge does not admit of traditional reductive analysis". They argue that knowledge is not reducible to other epistemic states, rather, it is in itself the basic building block of other epistemic states. But the idea of justification being an important epistemic phenomenon has already been imprinted in philosophical literature. See Kelp, Christoph. 2023. "How to be a capacitist." *Synthese* 201: 169.

4. Ben Kotzee holds that "[c]uriosity, open-mindedness and intellectual sensitivity are all epistemic virtues—they are all characteristics of a good knower inasmuch as they are conducive to that person's knowing well." See Kotzee, Ben. 2013. "Introduction: Education, Social Epistemology and Virtue Epistemology." *Journal of Philosophy of Education* 47 (2): 157-167.

5. It is important to note that in a later work, Harvey Siegel, together with Ben Kotzee and J. Adam Carter, endorse the cultivation of intellectual virtue, in addition to critical thinking, in the epistemic pursuits of education. See Kotzee, Ben, Carter, J. Adam and Harvey Siegel. 2021. "Educating for Intellectual Virtue: A Critique from Action Guidance." *Episteme* 18 (2): 177-199.

6. The year 1963 marks a pivotal period in analytic epistemology, with the publication of Edmund Gettier's critique of the traditional analysis of propositional knowledge.

REFERENCES

Amineh, Roya Jafari and Hanieh Davatgari Asl. 2015. "Review of Constructivism and Social Constructivism." *Journal of Social Sciences, Literature and Languages* 1 (1): 9-16.

Baehr, Jason. 2013. "Educating for Intellectual Virtues: From Theory to Practice." *Journal of Philosophy of Education* 47 (2): 248-262.

Battaly, Heather. 2014. "Intellectual Virtues." In *The Handbook of Virtue Ethics*, edited by Stan van Hooft, 177–187. Stocksfield: Acumen Publishing.

Battaly, Heather. 2016. "Responsibilist Virtues in Reliabilist Classrooms." In *Intellectual Virtues and Education: Essays in Applied Virtue Epistemology,* edited by Jason Baehr, 163-183. Oxon: Routledge.

Bringsjord, Selmer and Naveen Sundar Govindarajulu. 2018. 2018. "Artificial Intelligence." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta and Uri Nodelman. https://plato.stanford.edu/entries/artificial-intelligence/ (accessed October 22, 2025).

Buchanan, Tom. 1998. "Using the World Wide Web for Formative Assessment." *Journal of Educational Technology Systems* 27 (1): 71-79.

Carter, J. Adam and Emma C. Gordon. 2014. "Objectual Understanding and the Value Problem." *American Philosophical Quarterly* 51 (1): 1-13.

Chalmers, David J. 2010. "The Singularity: A Philosophical Analysis." *Journal of Consciousness Studies* 17 (9-10): 7-65.

Cochran-Smith, Marilyn and Susan L. Lytle. 1999. "Relationships of Knowledge and Practice: Teacher Learning in Communities." *Review of Research in Education* 24 (1): 249-305.

Corbett, Albert T. and Akshat Bhatnagar. 1997. "Student Modeling in the ACT Programming Tutor: Adjusting a Procedural Learning Model with Declarative Knowledge." In *User Modeling: Proceedings of the Sixth International Conference on User Modeling UM97,* edited Anthony Jameson, Cecile Paris and Carlo Tasso (Eds.), 243-254. Vienna: Springer-Verlag Wein GmbH.

Corbett, Albert. 2000. "Cognitive Mastery Learning in the ACT Programming Tutor." In *Adaptive User Interfaces: AAAI Technical Report SS-00-01*, edited by Seth Rogers and Wayne Iba, 31-36. California: AAAI Press. Available at https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=a3fc13be16d301136538b8a4c947fb5cb494aebc. (accessed October 15, 2025.)

Fungo, Joyce Estelle and Mark Anthony Dacela. 2022. "Objectual Understanding as the Primary Epistemic Aim of Education." *Kritike* 16 (1): 96-116.

Gettier, Edmund L. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23 (6): 121-123.

Goertzel, Ben. 2014. "Artificial General Intelligence: Concept, State of the Art, and Future Prospects." *Journal of Artificial General Intelligence* 5 (1): 1-46.

Goldman, Alvin I. 1999. *Knowledge in a Social World*. Oxford: Clarendon Press.

Greco, John. 2021. "Pritchard's Case for Veritism." *Epistemology & Philosophy of Science (Special Issue on Virtue Epistemology)* 58 (4): 46-53.

Hoc, Jean-Michel. 2000. "From Human-machine Interaction to Human-machine Cooperation." *Ergonomics* 43 (7): 833-843.

Hoover, Wesley A. 1996. "The Practice Implications of Constructivism." *SEDL Letter* 9, no. 3. https://sedl.org/pubs/sedletter/v09n03/practice.html (accessed October 9, 2025).

Hwang, Gwo-Jen, Xie, Haoran, Wah, Benjamin W. and Dragan Gašević. 2020. "Vision, Challenges, Roles and Research Issues of Artificial Intelligence in Education." *Computers and Education: Artificial Intelligence* 1 (100001): 1-5.

Käser, Tanja, Klingler, Severin, Schwing, Alexander G. and Markus Gross. 2017. "Dynamic Bayesian Networks for Student Modeling." *IEEE Transactions on Learning Technologies* 10 (4): 450-462.

Kelly, Peter. 2006. "What is Teacher Learning? A Socio-cultural Perspective." *Oxford Review of Education* 32 (4): 505-519.

Kelp, Christoph. 2023. "How to be a Capacitist." *Synthese* 201: 169.

Kohlberg, Lawrence and Rochelle Mayer. 1972. "Development as the Aim of Education." *Harvard Educational Review* 42 (4): 449-496.

Kotzee, Ben. 2013. "Introduction: Education, Social Epistemology and Virtue Epistemology." *Journal of Philosophy of Education* 47 (2): 157-167.

Kotzee, Ben, Carter, J. Adam and Harvey Siegel. 2021. "Educating for Intellectual Virtue: A Critique from Action Guidance." *Episteme* 18 (2): 177-199.

Kvanvig, Jonathan L. 2003. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.

Marabini, Alessia and Luca Moretti. 2020. "Goldman and Siegel on the Epistemic Aims of Education." *Journal of Philosophy of Education* 54 (3): 492-506.

Mascolo, Michael F. 2009. "Beyond Student-centered and Teacher-centered Pedagogy: Teaching and Learning as Guided Participation." *Pedagogy and the Human Sciences* 1 (1): 3-27.

Mason, Mark. 2008. "Complexity Theory and the Philosophy of Education." *Educational Philosophy and Theory* 40 (1): 4-18.

McArthur, David, Lewis, Matthew and Miriam Bishary. 2005. "The Roles of Artificial Intelligence in Education: Current Progress and Future Prospects." *Journal of Educational Technology* 1 (4): 42-80.

Melchior, Guido. 2023. "The Value of Knowledge and Other Epistemic Standings: A Case for Epistemic Pluralism." *Philosophia* 51: 1829-1847.

Metzler, Douglas P. and Cynthia J. Martincic. 1998. "Explanatory Mechanisms for Intelligent Tutoring Systems." In *ITS 1998: Intelligent Tutoring Systems (Lecture Notes in Computer Science, vol. 1452)*, edited by Barry P. Goettl, Henry M. Halff, Carol L. Redfield and Valerie J. Shute, 136-145. Heidelberg: Springer.

Morrison, Keith. 2008. "Educational Philosophy and the Challenge of Complexity Theory." *Educational Philosophy and Theory* 40 (1): 19-34.

Ouyang, Fan and Pengcheng Jiao. 2021. "Artificial Intelligence in Education: The Three Paradigms." *Computers and Education: Artificial Intelligence* 2 (100020): 1-6.

Powell, Katherine C. and Cody J. Kalina. 2009. "Cognitive and Social Constructivism: Developing Tools for an Effective Classroom." *Education* 130 (2): 241-250.

Pritchard, Duncan. 2013. "Epistemic Virtue and the Epistemology of Education." *Journal of Philosophy of Education* 47 (2): 236-247.

Riedl, Mark O. 2019. "Human-centered Artificial Intelligence and Machine Learning." *Human Behavior and Emerging Technologies: Perspectives from Technology Pioneers* 1 (1): 33-36.

Robertson, Emily. 2010. "The Epistemic Aims of Education." In *The Oxford Handbook of Philosophy of Education*, edited by Harvey Siegel, 11-34. New York: Oxford University Press.

Roll, Ido and Ruth Wylie. 2016. "Evolution and Revolution in Artificial Intelligence in Education." *International Journal of Artificial Intelligence in Education* 26: 582-599.

Shute, Valerie J., Gawlick-Grendell, Lisa A., Young, Robert K. and Clarke A. Burnham. 1996. "An Experiential System for Learning Probability: Stat Lady Description and Evaluation." *Instructional Science* 24: 25-46.

Siegel, Harvey. 2005. "Truth, Thinking, Testimony and Trust: Alvin Goldman on Epistemology and Education." *Philosophy and Phenomenological Research* 71 (2): 345-366.

Siegel, Harvey. 2010. "Critical Thinking." In *International Encyclopedia of Education, 3rd ed.,* edited by Penelope Peterson, Eva Baker and Barry McGaw, 141-145. Oxford: Elsevier Ltd.

Skinner, Burrhus Frederic. 1958. "Teaching Machines." *Science* 128 (3330): 969-977.

Stamper, John. 2006. "Automating the Generation of Production Rules for Intelligent Tutoring Systems." In *Proceedings of the 9th International Conference on Interactive Computer Aided Learning (ICL2006)*, 206–12. Villach: Kassel University Press. https://dev.stamper.org/publications/ICL2006-StamperFinal.pdf (accessed October 8, 2025).

Watson, Lani. 2016. "The Epistemology of Education." *Philosophy Compass* 11 (3): 146-159.

Yang, Stephen J.H., Ogata, Hiroaki, Matsui, Tatsunori and Nian-Shing Chen. 2021. "Human-centered Artificial Intelligence in Education: Seeing the Invisible through the Visible." *Computers and Education: Artificial Intelligence* 2 (100008): 1-5.

Zagzebski, Linda. 1996. *Virtues of the Mind: An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

Zagzebski, Linda. 2012. *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief*. Oxford: Oxford University Press.

* Australian Centre for Health Engagement, Evidence and Values, University of Wollongong, Australia