AI in learning: Preparing grounds for future learning

Hannele Niemi

Abstract
This special issue raises two thematic questions: (1) How will AI change learning in the future and what role will human beings play in the interaction with machine learning, and (2), What can we learn from the articles in this special issue for future research? These questions are reflected in the frame of the recent discussion of human and machine learning. AI for learning provides many applications and multimodal channels for supporting people in cognitive and non-cognitive task domains. The articles in this special issue evidence that agency, engagement, self-efficacy, and collaboration are needed in learning and working with intelligent tools and environments. The importance of social elements is also clear in the articles. The articles also point out that the teacher’s role in digital pedagogy primarily involves facilitating and coaching. AI in learning has a high potential, but it also has many limitations. Many worries are linked with ethical issues, such as biases in algorithms, privacy, transparency, and data ownership. This special issue also highlights the concepts of explainability and explicability in the context of human learning. We need much more research and research-based discussion for making AI more trustworthy for users in learning environments and to prevent misconceptions.

Keywords
artificial intelligence, human learning, machine learning, deep learning, explainability and explicability

AI is changing the learning landscape
We do not have a single common definition of artificial intelligence (AI), but we have a common understanding that AI is changing the world. It impacts societies, organizations, and work, and it is becoming more and more part of everyday life. We see AI-based technologies in different sectors of the economy and society, such as transportation, home, and service robotics, healthcare, education, public safety and security, employment, the workplace, and entertainment (e.g., Stone et al., 2016). These scenarios can be seen globally, for example, in China, the USA, and the European Union strategic plans. In China, a discussion paper from the McKinsey Global Institute (2017), originally presented at the 2017 China Development Forum, explored AI’s potential to fuel China’s productivity and growth—and to disrupt the nation’s workforce. It forecasts that AI technologies have far-reaching potential to improve healthcare, the environment, security, and education. AI, or the idea that computer systems can perform functions typically associated with the human mind, has gone from futuristic speculation to present-day reality. In the USA, the National AI Initiative Act of 2020 (House of Representatives, 2020) became law on January 1, 2021, providing for a coordinated program across the entire federal government to accelerate AI research and application for the nation’s economic prosperity and national security. The aim is to use trustworthy AI in the public and private sectors and prepare the present and future US workforce for the integration of AI systems across all sectors of the economy and society. Respectively, the European Commission’s (2020) white paper on AI sets strategic plans for how European countries will use AI in different sectors of society. It also includes a strong ethical stance and states. The European Union (2020)
summarizes “The European approach for AI aims to promote Europe’s innovation capacity in the area of AI while supporting the development and uptake of ethical and trustworthy AI across the EU economy. AI should work for people and be a force for good in society” (p. 8).

Even though we do not have any common definition of AI, the referred documents evidence that AI is strongly impacting our lives. In spite of their differences, the common visions have the following elements: AI means computational intelligence. Intelligent machines can examine data, make inferences, and then act by themselves (Roschelle et al., 2020). Many definitions also refer to machines as learning entities because they can adapt themselves to new tasks and make inferences based on interactions with other data-providing entities as well as with human beings.

In education and learning, applying AI to education is not new. The history of AI in learning research and educational applications goes back over 50 years (Minsky & Papert, 1968). In 2016, the researchers’ panel report summarized (Stone et al., 2016) that there have been several common AI-related themes worldwide over the past several years, such as teaching robots, intelligent tutoring systems (ITS), online learning, and learning analytics. The panel also indicated that massive open online courses (MOOCs) and other models of online education have been widely offered and used at all levels of educational systems, with sophisticated learning management systems that incorporate synchronous as well as asynchronous education and adaptive learning tools. In many studies, big data, learning analytics, and data mining techniques have been major tools for personalized learning as well as assessments (e.g., Baker & Inventado, 2014; Fischer et al. 2020).

However, recent technological rapid developments and new approaches to computing have set education and learning in a completely new context. The panel of 22 invited experts in learning sciences, education, and computing (Roschelle et al., 2020, p. 8) assessed the future and new designs of AI in learning and education: “These design concepts expand beyond familiar ideas of technology supporting ‘personalized,’ ‘adaptive,’ or ‘blended’ learning. The conventional metaphors may continue to be useful, but they also may limit how we envision futures of AI in learning.”

The aim of this commentary is to analyze what AI means in education and learning today and in the future. For this purpose, I have set two questions:

1. How will AI change learning in the future, and what role do human beings play in the interaction with machine learning?

2. What can we learn from the articles in this special issue for future research?

What is changing?

In the last 10 years, AI has taken big steps with a new method of computing and advanced technology for using and integrating multimodal data. The capacity of computational machines has grown to accommodate the analysis of a huge amount of data, which can be segmentalized or visualized for users to support, assess, or collect even weak signals to prevent dropping out from learning and education. Machine learning goes much beyond what was earlier possible with tracing users’ learning paths through keyboard strokes or eye movements in learning analytics. Recent developments in programming commonly use deep learning methods that imitate human brains (Pouyanfar et al., 2018; Shrestham & Mahmood, 2019). The starting point for programming is to build a computer program that resembles the neural networks of the human brain to consider the multiple connections between neurons and their networks. The program requires the analysis of a large amount of data by the machine. There can be several levels of networks, even dozens. The machine also constantly learns about the interactions that take place with the user. There may be several layers of neural networks, and the machine must be able to segmentalize, categorize, and visualize a large number of different data sets. The nature of the data that machines use can be multimodal. It can consist not only of text-based responses but also speech, facial expressions, gestures, and biological or physiological information, such as brain electrical curves, heart rate, or stress level measurements. With each use, the machine also learns something new. Dealing with a robot may irritate us, particularly when it does not understand our questions, but through our questions, the robot also learns.

In their report, the research panel mentioned earlier (Roschelle et al., 2020) describes how deep learning algorithms perform feature extraction in an automated way, which allows researchers to extract discriminative features with minimal domain knowledge and human effort. They summarize that these algorithms include a layered architecture of data representation, where the high-level features can be extracted from the last layers of the networks while the low-level features are extracted from the lower layers. These kinds of architectures were originally inspired by AI simulating its process of the key sensorial areas in the human brain. Our brains can automatically extract data representations from different scenes. The input is the scene information received
from the eyes, while the output is the classified objects. This highlights the major advantage of deep learning—that is, it mimics how the human brain works.

The panelists (Roschelle et al., 2020) describe recent developments in the multimodality of data, which goes beyond observing what students type on a computer or how they answer questions. Newer research-based systems can listen to recordings or watch videos of classrooms, finding events that are significant for learning outcomes (e.g., Suresh et al., 2019). Automated essay scoring is another long-standing application of AI (Page, 2003). It is also now rapidly expanding to include assistive systems for peer grading, student collaboration, and other educational applications of AI-based learning. New AI technologies, along with other emerging technologies, can produce learning innovations that include rigorous performance assessment, virtual reality, voice-based systems, gesture-based systems, social and educational robots, collaborative learning, mobile learning, and more.

In many countries, school leaders are also seeking to understand how new AI capabilities could become a reality in education and what should be taken into account in future planning (Wong, 2018; Roschelle et al. 2020). The following themes are in discussion:

- Perception via multiple sensors and the ability to recognize complex sets of features (e.g., use of cameras and motion detectors to recognize particular faces entering a building)
- Representation and reasoning, building models of people and their behaviors, and making inferences based on those models about what might happen next
- Learning and discovering meaningful patterns in large amounts of data
- Natural interaction (e.g., interacting through speech or gestures)
- Societal impact, leveraging infrastructures to do all of the above at a massive scale and in ways that directly affect people’s lives

Machines are learning, but how can we combine human learning with machine learning? Human learning has been investigated empirically for almost 200 years. Machine learning has been an object of scientific research for the past 30–50 years, depending on what is considered the origin of computer-based machine learning. How can these two research areas be combined? Cross-disciplinary research is still at an initial stage and needs much further investigation. We now have two learners: a human and a machine. How do these two entities work together? In many AI intelligent applications, both learners are also teaching each other. So far, we have very little understanding of what happens in a human–machine interaction.

In educational studies in the early 20th century, the prevailing theory of learning was a behaviorist theory derived from animal experiments. Human beings were regarded to have learned new things and skills when receiving reinforcement for the correct answer or action. This was also the theoretical basis of the first so-called teaching machines that gave a task or series of tasks to which learners reacted by giving their responses. Since the middle of the last century, behaviorism has provoked a lot of criticism among researchers who emphasized that learning takes place in a social and cultural context and that the environment has a major impact on what is learned (e.g., Vygotsky, 1978; Cole, 1991; Kozulin & Pressieen, 1995). As a promotor of social-cultural theory, Vygotsky emphasized that learning is a constant interaction between the learner, the material being taught, human beings, and the artifacts, for example, physical or psychological tools. Similarly, Jerome Bruner (1973) strongly pointed out that human beings learn not only facts but also meaningful entities in which a human being constructs representations and models that consist of meaningfully connected elements. Since the middle of the last century, there has been a powerful breakthrough in constructivist learning theories, in which humans are considered active creators and builders of knowledge (Greeno, 1996; Bransford et al., 2006). This was also the time when the concept of deep learning emerged in human learning, as opposed to surface learning, in which the learner did not understand or control learning units and the relationships of parts of knowledge (Marton & Säljö, 1976; Beattie, 1997; Biggs, 1988). The trend toward the structures of knowledge was subsequently reinforced as socio-constructivism, in which the creation of knowledge and skills is also considered a social process. People create knowledge together, which is also necessary for solving problems.

Current studies on learning have shown that there is a strong social dimension to learning. This was clearly observed when teaching moved online during the COVID-19 pandemic. Although all the information was available, teachers were able to organize teaching quite well, and the technological environment also worked well; however, the biggest criticism was for the lack of genuine social interaction (e.g., Niemi & Kousa, 2021). Both students and teachers missed being able to see other people with their expressions, gestures, and meet in a variety of real-world situations, from classrooms to hallways, dining, and cafes. Central to learning is the interaction that enables a shared experience. Learning has increasingly been regarded as embedded within a social context and framework. Social perspective theories have been variously called
social constructivism, sociocultural perspective, sociohistorical theory, and sociocultural-historical psychology. Although social perspective theorists’ views are diverse, each theorist posits that learning occurs through the mediation of social interaction (Niemi, 2009). Knowledge is not an individual possession but is socially shared and emerges from participation in social activities (Cole, 1991; Bransford et al., 2006).

The concept of metacognition has also influenced the concept of learning (e.g. Winne, 1996; Boekaerts, 1997; Biggs, 1988). Learning is an active process in which learners regulate and monitor their learning processes, and it depends on how self-regulated they are and how they use resources that are available. Self-regulated learners have a large arsenal of cognitive and metacognitive strategies that they readily deploy when necessary to accomplish academic tasks (Niemi, 2009). Further, self-regulated learners have adaptive learning goals and are persistent in their efforts to reach these goals (Schunk & Zimmerman, 1994). This leads to conceptualizing co-regulated learning in which learners set together aims for their learning and monitor their group processes. The learners’ active role is also linked with the concept of self-efficacy introduced by Bandura (1986). Its importance as an integral component of human agency has been confirmed in several studies on learning and learning outcomes, both in academic and non-academic tasks. (See Niemi & Niu, 2021, this issue.)

In recent years, biological and neurological studies of learning have brought new opportunities to understand what is happening in the human brain and body in the context of learning (Koizumi, 2003, 2004). Although the research is still in its initial stages in many ways, it is already known that a huge number of complex processes take place in the brain between neural networks, and neurons that form networks. Learning encompasses continuous and active interactions within and between neural networks. Learning is a holistic process in which the body, mind, and emotions are involved.

We can summarize that human learning is based on the following assumptions:

- Learning is an active process in which a human being is involved as a unique person.
- Learning is a social process, and learners are part of the social and cultural environment and context of his or her life.
- Learning happens in interactions with other people and artifacts that enhance human development.
- Human beings have agency in learning, and they monitor their learning through metacognitive strategies that include self- and co-regulative strategies. Motivational strategies, including engagement, and attributions of own learning, such as self-efficacy, have a strong impact on learning processes and products.
- Human learning is based on knowledge constructions that create meaningful entities.

New potential capacities of AI, such as the acquisition of multichannel data (voice, image, speech, text) and new sensor technologies, create extensive opportunities for understanding human learning and can open new ways to support human learning. However, interaction with machine learning requires much more research. We need much more understanding of what promotes and hinders human learning with machines. Can human beings have agency in AI-based learning and how can they monitor and attribute their own learning with machines? Can machine learning lead to a human learner’s meaningful knowledge construction? How can social elements be combined when learning with AI-supported environments?

What can we learn from the studies in this special issue?

The theme of this special issue is AI in learning, and the call invited researchers to submit articles that focused on learning in digital environments and with intelligent digital tools. We accepted seven articles with the criteria that they have a direct connection with AI and machine learning or that they can help to understand learning in digital-based learning environments for future AI-based tools. The seven articles were accepted and their major inputs will be introduced in the following summary. The first four articles of this special issue focus on student learning in intelligent digital environments. Three other studies are more related to teachers’ work and tasks when students are learning in digital environments.

The first article, “Evaluation of a practice system: Supporting distributed practice for novice programming students” (Li, B. et al., 2021) states that programming is an important skill in the 21st century, but it is difficult for novices to learn. By novice, they mean college students who are not engineers or computing science students. Programming contains many knowledge components, and its application scenarios feature great comprehensiveness and abstraction. Students sometimes fail to understand the most fundamental concepts and are unable to produce the most basic programs. The authors developed a mobile platform called Daily Quiz, which incorporated distributed practice, and the experimental group was encouraged to practice every three days. The control group practiced every seven days. The results showed that this simple manipulation
significantly improved the experimental group’s performance on the final exams. The experimental group of students achieved a higher rate of first-check correctness and tended to be more engaged in academic social interactions.

The second article, “Development and validation of computational thinking assessment of Chinese elementary school students” (Li, Y. et al., 2021), is related to computational thinking (CT). Its definition, teaching, and evaluation have been discussed by various scholars (e.g., Grover & Pea, 2018; Hsu et al., 2018; Nouri et al., 2019). Wing (2006) emphasized that CT is one of the daily life skills that everyone needs, rather than just being a programming skill used only by computer scientists. CT describes the processes and methods used to operate a system. The article focuses on computational thinking (CT) in school education. The authors develop a psychometrically validated assessment of CT literacy for children in Chinese elementary schools. Items are constructed to reflect key aspects of CT, such as abstraction, algorithm thinking, decomposition, evaluation, and pattern recognition. The analysis of the reliability and validity of the CTA–CES scale is examined carefully. The aim is to use the tool in measuring the literacy of CT of Chinese children, and perhaps, apply it to children worldwide.

These two articles focused on the themes of how students learn programming or how computational thinking (CT) can be measured. These themes have become urgent with AI. Machine learning is based on programming and algorithms. Should people understand how machines learn and how much they must know about it? These have become important questions in curriculum development in several countries. CT is also mentioned as future competences, often called 21st-century competences. There are many reasons for these discussions: future technology and its AI applications are part of our everyday life. Almost in all professions, advanced digital technology and AI applications will be used. Therefore, many colleges have added programming courses to their study programs, even though students are not becoming engineers or computing scientists.

Both of these studies provide evidence that we are at the beginning of combining two learners: a human learner and a machine learner. Human learners need to understand machine learning, and machines should understand human learning. Even though programming is not a new research area, it seems that we need new methods to support learners who are not professionals in this domain and who do not have basic knowledge of programming. We also need more understanding of what CT is and how to assess students’ learning processes and products when they learn CT.

The third study is related to student’s learning in a digital environment: “Digital storytelling enhancing Chinese primary school students’ self-efficacy in mathematics learning” (Niemi & Niu, 2021). The project was based on the assumption that learning is a socially and culturally related process that happens when interactions occur among learners, material tools, psychological tools, and other people (Vygotsky, 1978). In the study, interactions took place among students and between learners and mobile devices. The students created video stories on geometrical themes. The human–device interaction was a continuous and iterative process. The students used mobile phones and tablets with advanced video technology. They shot, edited, and modified their stories with different kinds of images and effects. Technology was intelligent, but the learners had agency in the human–machine interaction. The aim of this study was to uncover how digital storytelling advances students’ self-efficacy in mathematics learning. The teachers did not teach content but facilitated students’ work. The students self-assessed that they had become more confident that they could learn mathematics and understand what they had learned. They shared that they were very engaged in creating digital stories. The essential features were that the students worked collaboratively, they had strong agency in their learning, and they enjoyed working hard with their videos. We could conclude that in learning with intelligent tools, learners still need a sense of agency in learning, and they need to experience that learning is meaningful knowledge construction.

The fourth study is linked to both students and teachers. It provides an architecture of machine learning in the article “PLEA: A social robot with teaching and interacting capabilities” (Stipancic et al., 2021). The authors designed a learning environment in which a social robot head plays the role of a teaching assistant interacting with university undergraduates. PLEA facilitates student–teacher interaction. The robot could be autonomous or controlled remotely. It is based on natural interaction that is highly contextual and where participants analyze different inputs, such as previously memorized or currently sensed information, twisted, or guessed facts, and so on. The robot is capable of reasoning about human nonverbal communication signals and makes use of this capability based on the principles of multimodal interaction. The theoretical basis is provided by studies of human communication in psycholinguistics and social psychology. For the purposes of this study, three distinct sources of social signals were used: face emotion recognition, level of loudness, and intensity of body movements. Teachers can benefit from this information by adapting a presentation style and achieving better rapport with the student. Future studies will show the ways and extent
to which a cognitive robot can be truly effective in technology-enhanced learning.

The last three articles are focused on teachers’ work in digital environments or with AI-related tools. The expert panel (Roschelle et al. 2020) proposed that in the future, teachers’ role will be the orchestration of AI and other digital tools and environments. The future scenarios of AI in teaching and learning suggest that AI can be used as a toolkit to enable teachers and students to use different kinds of services, and to combine different intelligent tools that do not even exist today but will be developed in the future. The articles offered examples of the kinds of support teachers can receive from AI and how they can facilitate students’ learning in digital environments.

The fifth article in this special issue is “A dialogue system for identifying need deficiencies in moral education” (Chen et al., 2021). It is based on how teachers can support moral education by analyzing students’ problem behaviors and identifying their underlying need deficiencies. The authors first defined a theoretical framework to summarize all the factors relevant to students’ problem behaviors and need deficiencies. Thereafter, they developed a task-oriented dialogue system that can properly inquire about different aspects of students’ information and automatically infer their need deficiencies. They conducted comprehensive experiments to evaluate the system’s performance with real-life cases. The results show that the built dialogue system could effectively serve as a diagnostic tool to identify students’ deficiencies and help teachers. Through multturn dialogue, a task-oriented dialogue system can acquire the necessary information and complete the task automatically. The natural language interaction significantly improves service usability. The authors also reflect on the fact that AI-driven systems building and operating such intelligent agents would continuously collect a large amount of private user data and obtain sensitive analytical results. Hence, in practice, privacy-preserving techniques and policies are crucial to protecting students’ privacy and avoiding unnecessary surveillance.

The sixth study, “Designing a preliminary model of coaching pedagogy for synchronous collaborative online learning” (Timonen & Ruokamo, 2021), relates to teachers’ role in online learning. The study aimed to determine the kinds of synchronous collaborative online coaching pedagogy models that have been used in previous research. The methods comprised a systematic literature review and qualitative data and theory-driven content analysis of peer-reviewed articles spanning 2014–2018. The results identify several pedagogical frameworks for synchronous collaborative online learning: for example, the community of inquiry framework, including social, cognitive, and teaching presence; social presence in conjunction with the media synchronicity theory or the broaden-and-build theory; or the 4E learning cycle model (engagement, exploration, explanation, and extension). The preliminary results also indicate a scarcity of research on synchronous coaching pedagogy in online education. The authors constructed a preliminary pedagogical model for a coaching pedagogy for synchronous collaborative online learning (CPSCOL). The CPSCOL model focuses on verbal, textual, and collaborative content-sharing and learning, and also uses videos, visual, and nonverbal cues in online pedagogy. The aim of the CPSCOL principles of practice is to support practical implementations in which important elements are groups for peer learning and co-constructing problem-based learning and coaching processes to support cognitive goals and create social cohesion. The aim is to strengthen learner–learner dialogue and reflection through collaborative methods and reduce social distance via tools and methods of synchronous environments. The model emphasizes emotional engagement and human touch with the help of breakout groups and strengthens the online presence of learners through connections. The aim of group coaching is to foster internal agency and uphold the group’s common targets.

The seventh article, “Conceptualizing dimensions and a model for digital pedagogy” (Väätäinen & Ruokamo, 2021), is related to the question of what digital pedagogy is. The study is also based on literature review. The study conceptualized a model for digital pedagogy to provide tools for digital environments in teaching. The model of digital pedagogy is discussed in terms of three dimensions: (1) pedagogical orientation, (2) pedagogical practices, and (3) digital pedagogical competencies. The study examined how these dimensions are presented in the current research literature. The researchers reviewed articles published in the years 2014 to 2019. The findings suggest that, first, in many cases, pedagogical orientation is labeled socio-constructivist and student-centered. Second, pedagogical practices are the methods used to promote students’ learning, and they involve, for example, collaboration and social knowledge construction. Lastly, teachers’ success in blending digital technologies into their teaching is improved by high self-efficacy and strong peer-collaboration skills. Based on the literature analysis, digital pedagogy includes more than just the teacher’s perspective on teaching and learning; it must also include the students’ perspective. The teacher’s role is to work as a facilitator who uses student-centered teaching approaches, makes it possible for students to control their own learning processes, and encourages students in collaborative learning. Student engagement, problem-based exercises, collaboration, and social knowledge construction are essential in
digital pedagogical practices that require the teacher to possess several skills or competencies in creating digital environments.

All seven of the articles in this special issue give evidence that we have many concepts and perspectives to consider in terms of how AI will change learning and education. Some themes are focused on teachers' work and pedagogy, some delve more into students' learning processes with and for intelligent tools, and some research areas have direct links with developing new tools for interactive AI applications. All the articles provide pieces of the big change that AI will bring to learning and education. Learners' can be supported in different ways with multimodal intelligent tools and environments, and teachers can also gain from the use of AI-based interactive and dialogical tools. The articles suggest that we can help learners to understand computational thinking and that we can support non-engineering college students in learning programming by effective pedagogical arrangements. AI-based learning happens in interaction with machines and learners, and future workers need at least some understanding of how machines are learning. The articles also provide evidence that agency, engagement, self-efficacy, and collaboration are needed in learning and working with intelligent tools and environments. These elements are not new findings in learning research. However, their roles in AI-based learning are more important than in earlier environments because so many options will be available to support personal learning, but it also depends on how a learner uses them. The importance of social elements in learning emerged in almost all the articles. This also pertained to the teacher's role. The articles pointed out that the teacher's role in digital pedagogy primarily involves facilitating and coaching.

Future scenarios

The changes in learning landscapes will be enormous because of the huge potential of new AI technology and its future scenarios. We now have two learners—a human and a machine—and both should learn and act intelligently and can be described metaphorically as deep learners. Human learning connotes meaningful knowledge creation as opposed to surface learning, and deep learning of machines entails high intensity and many levels of different networks in programming. The opportunities to support learning are huge, but AI should be integrated with pedagogy and the needs of human learning. An expert panel (Roschelle et al., 2020, pp. 19–20) has discussed how AI could support learning “in terms of orchestrating complex learning activities with multiple people and resources, augmenting human abilities in learning contexts, expanding naturalistic interactions among learners and with artificial agents, broadening the competencies that can be assessed, and revealing learning connections that are not easily visible.” They stressed that these approaches go beyond familiar design concepts for individualized, personalized, or adaptive learning. To bring these approaches to life, the experts suggested that the panel make seven recommendations for research priorities:

1. Investigate AI designs for an expanded range of learning scenarios
2. Develop AI systems that assist teachers and improve teaching
3. Intensify and expand research on AI for assessment of learning
4. Accelerate development of human-centered or responsible AI
5. Develop stronger policies for ethics and equity
6. Inform and involve educational policymakers and practitioners
7. Strengthen the overall AI and education ecosystem

Many AI implementations are related to learning, but are not directly connected to cognitive or academic tasks, or they occur in environments that are not schools or other educational institutions. Learning is a broad concept and also covers non-cognitive mental processes. Often, cognitive and non-cognitive elements are integrated, and AI can even help to connect cognitive and non-cognitive processes in learning. Typical examples can be found in sports, art, and mental health services. In sports, AI applications have already been used for many years to measure several physical and mental indicators and to support people in learning more about their physical efforts. Kos et al. (2018) describe wearable devices that measure some physical or physiological quantity of an individual and that have become a part of daily life for many people. Microelectronic systems connect data from different sources and provide much information for improving performances (Grün et al. 2011). AI applications in sports are also widely used in training and coaching through real-time tracking systems.

For music and art productions, multimodal techniques can provide incredible opportunities to connect cognitive and non-cognitive learning processes (Cetic & She, 2021; Chen et al., 2020), and they can have a strong effect on the human mind and creativity. In medicine, AI has been widely applied (Wang & Preiminger, 2019). It has also been recently applied in mental health services, where human learning plays an important role. In addition to diagnosing problems, AI has also been applied to conversational treatments (Graham et al., 2019). However, the promises of AI for mental disorders have not yet reached full potential
because of the unmeasurable aspects of mental disorders, and the use of AI may lead to ethically and practically undesirable consequences (Uusitalo et al., 2021).

Today, AI for learning provides many applications and has broad potential in supporting people with cognitive and non-cognitive task domains. But it has also many limitations. Many worries are linked to ethical issues, such as biases in algorithms, privacy, transparency, and data ownership. Many scholars have expressed worries about the potential risks from the decisions that the machines make (Goebel et al., 2018; Floridi et al., 2018; Rai, 2020). So far, programs do not have mechanisms to explain their actions and behavior. Even though we understand the underlying mathematical scaffolding of current machine learning architectures, it is often impossible to get insight into the internal working of the models.

In education and learning, machines can make decisions that are related to, for example, entrance and access to the next educational step, certificate validations, performance automatic scoring, tutoring, and advice on assignments. Understanding the reasons for decisions are key issues in human behavior and communication and particularly in issues related to own personal learning. People want to have explanations for why something has happened or why some decision has been made. Human beings also reflect on their own learning. Metacognition is based on human learners’ capacity to assess, reflect on, and improve their own learning strategies. If that element is missing, AI-based support may be limited.

How much machines can explain their decisions has raised widespread discussion on transparency, often also expressed in terms of explainability and explicability (Coeckelbergh, 2020; Goebel et al., 2018; Robbins, 2019). They have become essential concepts and aims in ethical guidelines for AI. An AI system is not only expected to perform a certain task or make decisions but also to include a model with the ability to give a transparent report of why it took specific conclusions. Goebel et al. (2018) explain that the requirement of explainability is at least as old as early AI, but the recent and relatively rapid success of AI/machine learning solutions have made it more necessary and also more challenging than ever. New programming has arisen from neural network architectures. Even though we understand the underlying mathematical scaffolding of the current machine learning architectures, it is often impossible to get insight into the internal working of the models. Goebel et al. (2018, p. 2060) note that black boxes are implanted in deep learning architecture. We do not know how the system arrives at a particular decision relevant to a particular person. The system cannot explain or make transparent the decision-making in all its steps. Rai (2020) proposes a metaphor of glass boxes as an aim for computing architecture to make sense of decisions and add trustworthiness.

How machines can explain their decisions and be responsible for them is a contradictory area (Coeckelbergh, 2020; Goebel et al., 2018; Morley, 2020; Robbins, 2018). So many things are new and do not fit with traditional ethical considerations. Coeckelbergh (2020, p. 2062) proposes that what is important ethically speaking is not explainability as a feature of technical systems such as AI; the primary aim is explainability as answerability on the part of the human using and developing the AI. The technical “explainability,” i.e. what the AI system can “say” or “answer,” should be seen as something in the service of the more general ethical requirement of explainability and answerability on the part of the human agent who needs a sufficiently transparent system as a basis for the (potential) answers she gives to those affected by the technology.

The main responsible entity in decision-making falls to human beings—those who are designing and developing new applications and algorithms for machines and those who are using the applications.

The application of AI in multiple areas of the society has called forth a need to explore the possible threats and benefits related to the use of AI (Niemi, 2020). Floridi et al. (2018, p. 691) state that fear, ignorance, misplaced concerns, or excessive reactions may lead a society to underuse the full potential of AI. Goebel et al. (2018) forecasts that AI is becoming an increasingly ubiquitous co-pilot for human decision-making. We are in the beginning of co-learning and co-decision-making in human–machine interaction in terms of computing architecture and ethical challenges.

Floridi et al. (2018) reviewed several guidelines for ethically sustainable AI policy that lay the foundations for a “Good AI Society.” They present a synthesis of five ethical principles that should undergird its development and adoption and offer 20 concrete recommendations for national or supranational policymakers and other stakeholders. First, they set the following value basis for what beneficial AI use is about: enabling human self-realization without devaluing human abilities, enhancing human agency without removing human responsibility, increasing societal capabilities without reducing human control, and cultivating societal cohesion without eroding human self-determination. Based on these concepts, they confirm the four common principles that can be seen in most ethical guidelines for A, and they add a fifth one: explicability.
The five principles can be summarized as follows:

1. Beneficence: promoting well-being, preserving dignity, and sustaining the planet
2. Non-maleficence: privacy, security, and “capability caution.” Though “do only good” (beneficence) and “do no harm” (non-maleficence) seem logically equivalent, they represent distinct principles in both bioethics and the ethics of AI. The many potentially negative consequences of overusing or misusing AI technologies suggest the need for caution. Of particular concern is the prevention of infringements on personal privacy, which is linked to individuals’ access to and control over how personal data are used.
3. Autonomy: the power to decide (or whether to decide). The idea is that individuals have a right to make decisions for themselves about the treatment they do or do not receive. Thus, affirming the principle of autonomy in the context of AI means striking a balance between the decision-making power we retain for ourselves and that which we delegate to artificial agents.
4. Justice: promoting prosperity and preserving solidarity. These concepts are typically invoked in relation to the distribution of resources, such as new and experimental treatment options or simply the general availability, ensuring that the use of AI creates benefits that are shared (or at least shareable) and preventing the creation of new harms.
5. Explicability: enabling the other principles through intelligibility and accountability. This principle is expressed using different terms: “transparency,” “accountability,” “intelligibility,” and “understandable and interpretable.” Though described in different ways, each of these principles captures something seemingly novel about AI—that its workings are often invisible or unintelligible to all but (at best) the most expert observers.

For explicability, Floridi et al. (2018, p. 700) suggest that

for AI to promote and not constrain human autonomy, our “decision about who should decide” must be informed by knowledge of how AI would act instead of us; and for AI to be just, we must ensure that the technology—or, more accurately, the people and organisations developing and deploying it—are held accountable in the event of a negative outcome, which would require in turn some understanding of why this outcome arose. More broadly, we must negotiate the terms of the relationship between ourselves and this transformative technology, on grounds that are readily understandable to the proverbial person “on the street.”

AI in learning has the potential to lead to enormous innovations in learning, but for that, we need huge investments in research where human learning and intelligent machine learning are combined. We need more basic and applied research about AI using multimodal data. However, learning always happens in social and cultural contexts, and we also need more understanding of how teachers can integrate AI-based tools into their pedagogy in such a way that learners have agency and teachers have the capacity to orchestrate different digital tools, AI included. AI in learning is also a big societal issue and relates to how people’s privacy is ensured, how they understand what AI in learning is, and what the consequences are when wide multimodal data are gathered. We need much more research and research-based discussion for making AI more trustworthy for users and to prevent misconceptions. The issues of privacy and ownership are essential. Who owns the data in multimodal environments? Who can use the information and for what purposes? Who is responsible for decisions with AI tools and services? Who can explain what happens in human and machine learning and decision-making? These are big questions that should be urgently answered, and we need much cross-disciplinary research cooperation. We also have experience suggesting that technology companies that develop systems for learning need a more theoretical basis for learning and pedagogy. In this regard, much more cooperation between researchers, practitioners, and companies is needed.

We are now at a crossroads to destinations that will be led by AI in learning. This special issue opens doors, albeit slightly, to the future, by providing important examples of AI applications and implementations of digital learning and pedagogy. However, much more research is needed.

Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The author has received financial support from Business Finland for Co-innovation project AI in Learning (2020-2021).
References

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. Larusson, & B. White (eds.), Learning Analytics (pp. 61–75). New York: Springer. https://doi.org/10.1007/978-1-4614-3305-7_4.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. Journal of Social and Clinical Psychology, 4(3), 359–373. https://doi.org/10.1521/jscp.1986.4.3.359.

Bransford, J., Stevens, R., Schwartz, D., Meltzoff, A., Pea, R., Roschelle, J., Vye, N., Kuhl, P., Bell, P., Barron, B., Reeves, B., & Isabelli, N. (2006). Learning theories and education: Toward a decade of synergy. In P. A. Alexander & P. H. Winne (eds.), Handbook of educational psychology (pp. 209–244). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Beattie, V., Collins, B., & McInnes, B. (1997). Deep and surface learning: A simple or simplistic dichotomy? Accounting Education, 6(1), 1–12. https://doi.org/10.1080/09639289731587.

Biggs, J. (1988). The role of metacognition in enhancing learning. Australian Journal of Education, 32, 127–138. https://doi.org/10.1111/j.0044418803200201.

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. Learning and Instruction, 7(2), 161–186. http://dx.doi.org/10.1016/S0926-6314(96)00015-1.

Bruner, J. (1973). Beyond the information given. Studies in the psychology of knowing. New York: Norton. https://arch.ive.org/details/beyondinformation00brunrich/mode/2up.

Cetinic, E., & She, J. (2021). Understanding and creating art with AI: Review and outlook. New York: Cornell University Press. arXiv:2102.09109.

Chen, P., Lu, Y., Yu, S., Xu, Q., & Liu, J. (2021, this issue). A dialogue system for identifying need deficiencies in moral education. Journal of Pacific Rim Psychology, 15, 1–14. http://dx.doi.org/10.1108/jpp.2018.07.004.

Chen, Z., Chen, L., Zhao, Z., & Wang, Y. (2020). AI illustrator: Art illustration generation based on generative adversarial network. 2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC), Beijing, 2020. 55–159. doi:10.1109/ICIVC50857.2020.9177494.

Coeckelbergh, M. (2020). Artificial intelligence, responsibility attribution, and a relational justification of explainability. Science and Engineering Ethics, 26, 2051–2068. https://doi.org/10.1007/s11948-019-00146-8.

Cole, M. (1991). Conclusion. In L. B. Resnick, J. M. Levine, & S. D. Teasley (eds.), Perspectives on socially shared cognition (pp. 398–417). Washington, DC: American Psychological Association.

European Commission (2020). A European approach to artificial intelligence. Retrieved from https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence.

European Union (2020). European framework on ethical aspects of artificial intelligence, robotics and related technologies. Retrieved from https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654179/EPRS_STU(2020)654179_EN.pdf.

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. Review of Research in Education, 44(1), 130–160. https://doi.org/10.3102/00917322X2093304.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luette, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P. & Vayena E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28, 689–707. https://doi.org/10.1007/s11023-018-9482-5.

Goebel, R., Chander, A., Holzinger, K., Leece, F., Akata, Z., Stumpf, S., Kieseberg, P., & Holzinger, A. (2018). Explainable AI: The new 42? In A. Holzinger, P. Kieseberg, A. Tjoa, & E. Weippl (eds.), Machine learning and knowledge extraction. Lecture Notes in Computer Science, vol. 11015 (pp. 295–303). Germany, Hamburg: Springer. https://doi.org/10.1007/978-3-319-99740-7_21.

Graham, S., Depp, C., Lee, E.E., Nebeker, Tu, X., Kim, H.-C., & Jeste, D.V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. Current Psychiatry Reports, 21(116). https://doi.org/10.1007/s11920-019-1094-0.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (eds.), Handbook of educational psychology (pp. 15–46). Macmillan Library Reference USA. New York: Prentice Hall International.

Grover, S., & Pea, R. (2018). Computational thinking: A competency whose time has come. In S. Sentance, E. Barendsen, & C. Schulte (eds.), Computer science education: Perspectives on teaching and learning (pp. 19–38). London: Bloomsbury Academic.

Grün, T., Franke, N., Wolf, D., Witt, N., Eidloß, A. (2011) A real-time tracking system for football match and training analysis. In A. Heuberger, G. Elst, & R. Hanke (eds.), Microelectronic systems (pp. 199–212). Berlin: Springer. https://doi.org/10.1007/978-3-642-23071-4_19.

House of Representatives (2020). Report 116–617. Retrieved from https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf?page=s1210.

Hsu, T-C., Chang, S-C., & Hung, Y-T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. Computers & Education, 126, 296–310. https://doi.org/10.1016/j.compedu.2018.07.004.

Koizumi, H. (2003). Science of learning and education: An overview. Studies in the psychology of knowing. New York: Norton. https://arch.ive.org/details/beyondinformation00brunrich/mode/2up.
Koizumi, H. (2004). The concept of “developing the brain”: A new natural science for learning and education. *Brain and Development, 26*(7), 434–441. https://doi.org/10.1016/j.braindev.2003.09.011.

Kos, A., Wei, Y., Tomažič, S., & Umek, A. (2018). The role of science and technology in sport. *Procedia Computer Science, 129*, 489–495. https://doi.org/10.1016/j.procs.2018.03.029.

Kozulin, A., & Presseisen, B. Z. (1995). Mediated learning experience and psychological tools: Vygotsky’s and Feuerstein’s perspectives in a study of student learning. *Educational Psychologist, 30*(2), 67–75. http://dx.doi.org/10.1207/s15326985ep3002_3.

Li, B., Ning, F., Zhang, L., Yang, B., & Zhang, L. (2021, this issue). Development and validation of computational thinking assessment of Chinese elementary school students. *Journal of Pacific Rim Psychology, 15*, 1–22. http://dx.doi.org/10.1177/18344909211008264.

Li, Y., Xu, S., & Liu, J. (2021, this issue). Development and validation of computational thinking assessment of Chinese elementary school students. *Journal of Pacific Rim Psychology, 15*, 1–22. http://dx.doi.org/10.1177/18344909211008264.

McKinsey Global Institute (2017). Artificial intelligence: Implications for China, April 2017, discussion paper. Presented at the 2017 China development forum. Retrieved from https://www.mckinsey.com/.../media/mckinsey/featured%20insights/china/artificial%20intelligence%20implications%20for%20china/mgi-artificial-intelligence-implications-for-china.pdf.

Minsky, M. L., & Papert, S. A. (1968). *Perceptrons: An introduction to computational geometry*. Cambridge, MA: MIT Press.

Morley, J., Floridi, L., Kinsey, L., & Alhalal, A. (2020). From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practice. *Science and Engineering Ethics, 26*, 2141–2168. doi: 10.1007/s11948-019-00165-5.

Niemi, H. (2020). Artificial intelligence for the common good in educational ecosystems. In *Humanistic futures of learning: Perspectives from UNESCO Chairs and UNITWIN Networks* (pp. 148–152). Paris: Ediciones UNESCO. Retrieved from https://researchportal.helsinki.fi/en/publications/artificial-intelligence-for-the-common-good-in-educational-ecosys.

Niemi, H. (2009). Why from teaching to learning? *European Educational Research Journal, 8*(1), 1–17. https://doi.org/10.2304/eerj.2009.8.1.1.

Niemi, H., & Kousa, P. (2020). A case study of students’ and teachers’ perceptions in a Finnish high school during the COVID pandemic. *International Journal of Technology in Education and Science, 4*(4), 352–369. https://doi.org/10.46328/ijtes.v4i4.167.

Niemi, H., & Niu, S. J. (2021, this issue). Digital storytelling enhancing Chinese primary school students’ self-efficacy in mathematics learning. *Journal of Pacific Rim Psychology, 15*, 1–17. http://dx.doi.org/10.1177/18344909211991432.

Nouri, J., Zhang, L., Mannila, L., & Norén, E. (2019). Development of computational thinking, digital competence and 21st century skills when learning programming in K-9. *Education Inquiry, 11*(1), 1–17. https://doi.org/10.1080/20004508.2019.1627844.

Page, E. B. (2003). Project essay grade: PEG. In M. D. Shermis & J. Burstein (eds.), *Automated essay scoring: A cross-disciplinary perspective* (p. 43). Mahwah, NJ: Lawrence Erlbaum Associates.

Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., Shyu, M.-L., Chen, S.-C., & Iyengar, S. S. (2018). A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys, 51*(5), 1–36. https://doi.org/10.1145/3234150.

Rai, A. 2020. Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science, 48*, 137–141. https://doi.org/10.1007/s11747-019-00710-5.

Robbins S. (2019). A misdirected principle with a catch: Explicability for AI. *Minds and Machines, 29*, 495–514. https://doi.org/10.1007/s11023-019-09509-3.

Roschelle, J., Lester, J., & Fusco, J. (eds.) (2020). *AI and the future of learning: Expert panel report*. Digital Promise. Retrieved from https://ciresl.org/reports/ai-report.

Schunk, D., & Zimmerman, B. (1994). Self-regulation of learning and performance: Issues and educational applications. Hillsdale, NJ: Erlbaum.

Shrestham, A., & Mahmood, A. (2019). Review of deep learning algorithms and architectures. *IEEE Access, 7*, 53040–53065.

Stipancic, T., Koren, L., Korade, D., & Rosenberg, D. (2021, this issue). PLEA: A social robot with teaching and interacting capabilities. *Journal of Pacific Rim Psychology, 15*, 1–13. http://dx.doi.org/10.1177/18344909211037019.

Stone, P., & 16 Study Panel members (2016). *Artificial intelligence and life in 2030, one hundred year study on artificial intelligence report of the 2015*. Stanford, CA: Stanford University. Retrieved from https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100report10032016fnl_singles.pdf.

Suresh, A., Sumner, T., Jacobs, J., Foland, B., & Ward, W. (2019). Automating analysis and feedback to improve mathematics teachers’ classroom discourse. *Proceedings of the AAAI Conference on Artificial Intelligence, 33*(1), 9721–9728. https://doi.org/10.1609/aaai.v33i01.33019721.

Timonen, P., & Ruokamo, H. (2021, this issue). Designing a preliminary model of coaching pedagogy for synchronous collaborative online learning. *Journal of Pacific Rim Psychology, 15*, 1–22. http://dx.doi.org/10.1177/1834490921991430.

Uusitalo, S., Tuominen, J., & Arstila, V. (2021). Mapping out the philosophical questions of AI and clinical practice in diagnosing and treating mental disorders. *Journal of
Evaluation in Clinical Practice, 27(3), 478–484. https://doi.org/10.1111/jep.13485.

Väätäinen, J., & Ruokamo, H. (2021, this issue) Conceptualizing dimensions and a model for digital pedagogy. Journal of Pacific Rim Psychology, 15, 1–12. https://doi.org/10.1177/1834490921995395

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Wang, F., & Preininger, A. (2019). AI in health: State of the art, challenges, and future directions. Yearbook of medical informatics. Stuttgart: IMIA and Georg Thieme Verlag KG. https://www.thieme-connect.com/products/ejournals/html/10.1055/s-0039-1677908

Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33–35.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. Learning and Individual Differences, 8(4), 327–354. http://dx.doi.org/10.1016/S1041-6080(96)90022-9.

Wong, G. K. W., Ma, X., Dillenbourg, P., & Huan, J. (2018). Broadening artificial intelligence education in K-12: Where to start? ACM Inroads 11(1), 20–29.