This is the published version of a paper presented at *European Conference on the Impact of Artificial Intelligence and Robotics 2019 (ECIAIR 2019).*

Citation for the original published paper:

Teacher-supported AI or AI-supported teachers?  
https://doi.org/10.34190

N.B. When citing this work, cite the original published paper.

Permanent link to this version:  
http://urn.kb.se/resolve?urn=urn:nbn:se:miun:diva-37621
Teacher-Supported AI or AI-Supported Teachers?

Niklas Humble and Peter Mozelius  
Mid Sweden University, Östersund, Sweden  
niklas.humble@miun.se  
peter.mozelius@miun.se  
DOI: 10.34190/ECAIR.19.007

Abstract: Today, AI has a rapid dissemination and is becoming involved in many new areas. One of these areas is education, with AI in education (AIED) as an emerging research topic. When AIED will be applied in educational settings, would AI be implemented as a teacher support, or will the teachers support the AI systems? The aim of this study was to analyse contemporary research on AIED and discuss its possibilities and challenges. This study was carried out as a literature review, where found themes and patterns have been categorised in a thematic analysis. Important main keywords in the search where: human-compatible AI, teachers, education, human-supported AI, Al-supported humans, AI in education. Keywords have been combined with Boolean operators to find articles containing information that was important to answer the research question. In the deductive analysis found themes and patterns where grouped into the categories: teacher-supported AI, AI-supported teachers, teacher-compatible AI. Findings indicate that many studies lack a clear distinction between teacher-supported AI and Al-supported teachers. The recommendation is teacher-compatible AI, and that the combination of humans and AI is stronger than just one of them.

Keywords: AI in education, AIED, teacher-compatible AI, teacher-supported AI, AI-supported teachers

1. Introduction and aim

Artificial intelligence (AI) is a rapid growing field encompassing a number of subfields, such as learning and perception, proving mathematical theorems, playing chess, driving cars, writing poetry, and diagnosing diseases (Russel & Norvig 2016:1). There are many ways of defining AI, but what is central in many definitions is the ability to accomplish complex goals through human-like reasoning or acting (Tegmark 2017:50; Bostrom 2017:27; Russel & Norvig 2016:1-2). AI could also be understood as a further step in the human strive to understand intelligence through building intelligent entities (Russel & Norvig 2016:1).

A younger subfield of AI is artificial intelligence in education (AIED), which has expanded during the recent years (Koedinger & Corbett 2006; Heffernan & Heffernan 2014; Luckin et al 2016). AIED is, as AI, an interdisciplinary field containing psychology, linguistics, neuroscience, education, anthropology and sociology with the goal of being a powerful tool for education and providing a deeper understanding of how learning occurs (Self 1998; Luckin et al 2016).

In contemporary AIED research a number of potential applications of AI in education are suggested and discussed, which in extension could heavily affect the teachers, forcing them to take the step from “sage on the stage” to “guide on the side” (King 1993). The question arises: Will AIED support the teacher in teaching and learning activities or will the teacher support the AIED-system, or even be replaced by it?

The aim of this study was to analyse contemporary research on AIED and discuss its possibilities and challenges.

2. Extended Background

This section presents a brief history of artificial intelligence (AI) and artificial intelligence in education (AIED) together with some key aspects and concepts in the discussion of each field.

2.1 Artificial intelligence

The term artificial intelligence (AI) was coined in 1956 by John McCarthy (McCarthy et al 2006), but can be traced back to 1315 and Ramon Llulls Ars Magna where the idea that reasoning and thought process could be implemented artificially in an intelligent machine was outlined (Jensen 2017). However, the modern idea of AI was formalised by Alan Turing (Castelfranchi 2013), even though he did not use the term AI. Turing presented an idea of how intelligent computing machines could formalise intelligent reasoning and calculation (Turing 1937; Turing 1950).
A big part of the modern discussion of AI is the achieved, and potential, level of intelligence and its consequences on humanity, discussed by, among other, Nick Bostrom (2017) and Max Tegmark (2017). Some central concepts in this discussion are: weak AI, strong AI, narrow AI, and artificial general intelligence (AGI).

Considering the first two, weak and strong AI, John Searle (1990) explains the difference between the two as that strong AI is an AI that can perform in such a way that its performance cannot be distinguished by that of a human (the Turing test). While weak AI is a much more cautious view on AI where it is seen as computer models that can be useful for studying other fields, for example the weather, economics or the mind (Searle 1990). These two have recently been renamed to narrow AI and AGI, and further specified. Tegmark defines strong AI as AGI and AGI as the “Ability to accomplish any cognitive task at least as well as humans” and narrow intelligence as the “Ability to accomplish a narrow set of goals, e.g., play chess or drive a car” (2017:39).

2.2 Artificial intelligence in education

The field of AIED have since the early 1970s had an interest in investigating how AI techniques can be leverage to produce tools for personalised instructions that are tailored to the learners’ specific needs (Conati, Porayska-Pomsta & Mavrikis 2018). In the early days, the field of AIED could be dismissed as pedagogically narrow or naïve and technically bound, but has since then matured (Cumming & McDougall 2000).

One aspect of AIED is its potential to more cost efficiently provide each student with a personal tutor, rather than hiring a human tutor for every student (Koedinger & Corbett 2006). The aim of bringing AIED-systems to the same level of efficiency as a one-to-one human tutor have been a focused area for researcher since the earliest days of computers (VanLehn 2011). Koedinger and Corbett (2006) brings up five important aspects that a computer needs to be able to do to fully function as a tutor in the same way as a human can: 1) solve problems and reason through the use of domain knowledge, 2) have an understanding of typical learning trajectories, misconceptions and informal knowledge, 3) being able to follow the reasoning of a student step-by-step and understand when and where there is a lack in knowledge or understanding, 4) provide feedback, assistance and scaffolding when needed and in the right context, 5) individualise instructions in an ongoing assessment.

3. Theoretical assumptions

The theoretical assumptions in this study takes its starting point in ‘human-compatible AI’ and ‘the value alignment problem’. The value alignment problem is a discussion concerning the question of ensuring that the values of a potential strong AI or AGI is an alignment with our own (Fisac et al 2017; Hadfield-Menell et al 2016; Russell & Norvig 2016:1037; Wiener 1960). Human-compatible AI is a recent suggested development strategy for AI (Russell 2017) that potentially could solve the value alignment problem in that it proposes a possible co-existing relationship between human and AI. The idea is that AI is developed in such a way that it has objectives, but that those are only to maximise the realisation of human values which are not specified in the AI system but rather obtained, for example through observation of human behaviour (Russell 2017). This relationship standpoint could further be exemplified by the suggested move from a single-agent AI to a multi-agent AI, where the human is introduced as the second agent that sets the objective and values (Fisac et al 2017; Hadfield-Menell et al 2016).

With human-compatible AI as a midground, encompassing both the existing weak or narrow AI and potential strong AI or AGI, the other two possible solutions to the value alignment problem is based on, and mainly dependent on, one or the other. If AI remains relatively weak or narrow, lacking the cognitive ability to perform any critical tasks on its own, it could mainly function as a support for humans (Searle 1990) and the risk of giving the system explicit objectives of its own would probably be minimal. As noted by Baker (2016), this is generally the case for many AI systems in educational setting today. However, this is a relationship that can only be upheld as long as AI remains relatively dumb (Tegmark 2017:259), since it offers limited, or no, preparation for the potential evolvement to strong AI or AGI.

The last potential outcome of the value alignment problem is that no, or faulty, preparation is done in the development of AI and it reaches a human, or above, level of general intelligence. Some studies in the field of AIED indicates that AI systems are closing in, or even surpassing, humans in some aspects of their work to support students (Popenici & Kerr 2017; Kulik & Fletcher 2016; Johnson & Lester 2016; Rickel 2001). Now, where does this put humanity? It could reduce us to system supporters, as can already be seen today in the
use of MOOC courses (Roll & Wylie 2016). It could however also, in alignment with the concerns of AI systems taking over teachers’ jobs (Wogu et al 2018), put humanity in the inconvenient position of being unemployable due to economic gain and outperformance by intelligent machines (Tegmark 2017:123; Bostrom 2003).

4. Method and data collection

This study was conducted as a literature review to critically engage in the existing research of the studied field (Bryman 2016:6). Since the interests of the study is in contemporary research on AIED the literature search was limited to works published between 2015 and 2019, with the exceptions of backward-searches. The main keywords in the search where: human-compatible AI, teachers, education, human-supported AI, AI-supported humans, AI in education. These were further combined with Boolean operators in the search to find relevant articles, containing information important for the study’s aim. Google Scholar was used as the main search engine and a total of 20 articles was selected for further analysis.

A risk in literature reviews is the production of lists of findings and citations without a plot (Bem 1995), therefore a thematic analysis was used to systematically identify and organise meaningful and common themes in the literature search relevant to the aim of the study (Braun et al 2019; Bryman 2016:584).

However, what is common is not always relevant (Braun et al 2019), therefore deductive analysis was further used to group the findings in pre-decided categories (Rivas 2012). The pre-decided categories (teacher-supported AI, AI-supported teachers, and teacher-compatible AI) were based on the potential relationship outcomes between human and AI presented in the theorical assumptions.

Table 1: Selected articles

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Title</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Popplewell &amp; Harding</td>
<td>Impact of Moderation through the Distributed Virtual Enterprise Life Cycle.</td>
<td>Coventry University, School of Engineering.</td>
</tr>
<tr>
<td>2008</td>
<td>De Laat, Chamrada &amp; Weegerif</td>
<td>Facilitate the facilitator: Awareness tools to support the moderator to facilitate online discussions for networked learning.</td>
<td>Proceedings of the 6th international conference on networked learning (pp. 80-86).</td>
</tr>
<tr>
<td>2016</td>
<td>Luckin et al</td>
<td>Intelligence unleashed: An argument for AI in education.</td>
<td>Pearson</td>
</tr>
<tr>
<td>2016</td>
<td>Nye</td>
<td>Its, the end of the world as we know it: transitioning aied into a service-oriented ecosystem.</td>
<td>International Journal of Artificial Intelligence in Education, 26(2), 756-770.</td>
</tr>
</tbody>
</table>
5. Findings and discussions

The findings are presented in three categories: teacher-supported AI, AI-supported teachers, and teacher-compatible AI; followed by a general discussion. The common denominator for findings presented in teacher-supported AI is that they in some sense rely on strong AI or AGI to reach their full potential as described in the research. In AI-supported teachers the common denominator among the findings are that the AI-systems are not necessarily in need of intelligence to work as they are described. In teacher-compatible AI the common denominator is that the systems require some level of AI, narrow to intermediate AGI, for them to fully function as described in the research.

5.1 Teacher-supported AI

An example of teacher-supported AI can be found in the description and potential of intelligent learning environments (ILE) and intelligent tutoring systems (ITS). Researchers describes the strengths of ILE- and ITS-systems, in relation to a human tutor, in that it is always present, portable, and collaborative (Goksel & Bozkurt 2019; Nichols & Holmes 2018; Roll & Wylie 2016). Some studies indicate that ITS now is at the level of matching the tutoring of a human in regards of raising the test scores of students (Kulik & Fletcher 2016). ILE and ITS, as described here, also matches Sellars (2018) description of interactive and adoptive learning environments potential use in future education. Where the system is not only able to perform content and analytical oriented tasks, but also based the learnings affective aspects provide feedback and interactions with the students (Sellars 2018).

A second example of teacher-supported AI that is described as having the potential of, more or less, replacing multiple roles in education is intelligent virtual agents. These are described as having the potential of playing multiple roles, such as an artificial student on a similar cognitive level as the ones it is collaborating with (but with capabilities to introduce ideas on a higher level), an expert participant, or a participant that the students themselves can teach (Luckin et al 2016; Leelawong & Biswas 2008; Vizcaino 2005; Rickel 2001). A strength mentioned about the use of intelligent virtual agent in education is that research indicates that some students
might feel more comfortable training social skills, such as negotiating, with intelligent virtual agents (Gratch, DeVault & Lucas 2016). Intelligent virtual agents are still quite limited in their interactions with students and in need of further development (Johnson & Lester 2016; Rickel 2001). However, research indicates that intelligent virtual agents are becoming more emotionally and socially intelligent and are closing in on their potential as guides, mentors and teammates to students (Johnson & Lester 2016; Rickel 2001).

5.2 AI-supported teachers
An example of AI-supported teachers is the task oriented AIED systems or AIED as an ecosystem. The idea is that AIED needs to follow the recent trends in ubiquitous and service-oriented computing with multiple distributed systems and move away from the monolithic idea of a single AIED system (Nye 2016; Luckin et al 2016). Research suggest that there is great potential in using machine learning and edge computing in the combination of internet of things and smart cities in the context of education, for example to track and control students’ digital devices in the classroom (Pacheco et al 2018). A relating, and potential extended, use of this is AI analytics, and as mentioned by Nicholas and Holmes (2018) there is still a need for discussions and guidelines concerning the ethical implications of the use of AI in education, since much is still unclear and without a strong sense of directions for its use.

Another example of AI-supported teachers can be found in the idea of using AIED to amplify the intelligence of existing human (Pentland, Daggett & Hurley 2017). This can be said to solve an identified problem in many intelligent tutoring systems (ITS), that they simply are not that intelligent (Baker 2016). Baker (2016) classifies many of the existing tutoring systems as stupid tutoring systems as opposed to intelligent tutoring systems (ITS). This presents an alternative view on what ITS and AIED could be and Baker (2016) states that we do not need intelligent tutoring systems as long as we have intelligent human tutors. The ITS needs not to be intelligent but rather designed intelligently to amplify the human tutors already existing intelligence (Baker 2016).

5.3 Teacher-compatible AI
An example of teacher-compatible AI can be found in the potential use of digital assistants. In a recent study a discussion was made of using AI in education as a digital assistant in the post-digital classroom, to support the teacher in individualising the learning material and quantity training (Hrastinski et al 2019). This idea is supported in other research where the gain of using AIED or cobots (co-working robots) to free the teacher from routine task, keeping track on all the students, and possessing all the knowledge that might be relevant to the students are discussed (Goksel & Bozkurt 2019; Luckin et al 2016; Roll & Wylie 2016; Timms 2016). In the future, teachers might, like previously been done in the industry of manufacturing, have to possibility to program and reprogram educational cobots to suit their needs with the use of simple programming interfaces, such as block programming, that do not require advanced knowledge in programming (Nair, Kuhn & Hummel 2019).

Related to digital assistants is the idea of intelligent moderation in education. The strength of intelligent moderation is that it can identify conflicts, for example in decision-making, by combining different information and knowledge that it is being provided (Popplewell & Harding 2004). In educational setting intelligent moderation is described as playing its part in the context of large student numbers, where the vast amount of data produced by the students is too great for a single human teacher to process (Luckin et al 2016). AI techniques, such as text processing and machine learning, can be used by the intelligent moderation-system to summarise and analyse the data produced by the students. Which can later be reported back to help the teacher to better guide the students, for example by alerting when students gone off track or are in need of support. (Luckin et al 2016; De Laat, Chamrada & Wegerif 2008)

5.4 General discussion
All potential uses of AIED-systems presented above would have implications on the role of the teacher as it is today, which is also a concern that is mentioned in previous research (Hrastinski et al 2019). The systems presented under AI-supported teacher would probably not threaten to replace the job of the teacher, due to its limitations in intelligence, but could potentially decrease the need of teachers, tutors and administrative personnel due to its potential lower cost (Koedinger & Corbett 2006). However, there is also the question of how efficient these systems could be. Will they support the teacher or give the teacher more maintenance and administration (Roll & Wylie 2016)?
With the systems presented under teacher-supported AI we would have to move into a territory of AI that correlates with the description of strong AI or AGI (Tegmark 2017:39; Goertzel 2014) for them to fully operate as described. The question here would probably be to what degree teachers are needed to support the system, or if they are needed at all? This, as the potential decreasing need of education personnel, is a raising concern lifted in research on AIED (Wogu et al 2018; Popenici & Kerr 2017). Nicholas and Holmes (2018) suggests that an ethical framework must be developed and continuously updated for the work and use, and potential use, of AIED to reflect its possible capability and breadth in education. A notion also supported by Russell and Norvig, who declare that “Those who strive to develop AI have a responsibility to see that the impact of their work is a positive one” (2016:1051).

The findings presented under teacher-compatible AI could be said to be a midground between AI-supported teachers and teacher-supported AI in question of its need for intelligence. Ranging from somewhere between narrow AI to intermediate AGI, to operate as described. The systems presented here would, however, also incorporate both risks of previous categories: due to technical limitations being too narrow to bring anything essential to the collaboration or being that strong that it questions the need for a human teacher. Pentland, Dagget and Hurley (2017) presents a potential solution to this dilemma in that an AI-system must be able to visualise and explain the process in the system to achieve calibrated trust between the system and the human.

This would mean that the teacher, and other stakeholders, gets full insights in what the AI system really can do and on what ground it does it, and from there adopt the use of the system and the role of the teacher.

Regardless of the form AIED takes there are several existing and potential issues that needs to be addressed. How do we ensure that the system is representative and free from biases and personal agendas from the developers (Nichols & Holmes 2018; Popenici & Kerr 2017)? How do we balance standardisation and what is expected of a student with individualisation and customisation of the learning material by AIED (Walkington & Bernacki 2019; Wogu et al 2018)? A potential solution suggested in previous research is that teachers become co-designers of the activities in the classroom that are AIED-supported (Hrastinski et al 2019). This would further prevent the AIED-systems to be developed as black-boxes without transparency (Pentland, Daggett & Hurley 2017).

Finally, as stated by the former chess world champion Garry Kasparov “man plus machine is stronger that either” (Kasparov 2008:57). This is part of the discussions on ‘Advanced chess’ in the book, a discipline where human chess players assisted with computer chess software play against each other on a level that neither human grandmasters nor world leading computer chess programs are able to if not combined. Since then, new machine learning techniques have challenged the traditional heuristic-based chess programs. An example is when AlphaZero developed by the London based company Deep Mind defeated Stockfish in a series of chess games (Chess.com 2018).

However, as pointed out by the Stockfish developer Todd Romstad, Stockfish played with parameter settings and a time control that was unfavourable (Chess.com 2018), and that AlphaZero was run on much stronger hardware (Silver et al 2018). What would be interesting is if AlphaZero in the future accepts the invitation to play the yearly World Computer Chess Championship. A competition where all chess software must be run on the same hardware platform. Nevertheless, the idea of life imitating chess, Kasparov (2008:58) still seems reasonable and that human strategy and evaluation skills together with computer assistance are stronger than just one of them, and have the potential to reinvent many professions.

6. Conclusions

The conclusion of this study is that, even though contemporary research in AIED discuss different aspects and uses of AI-systems in education, it is often quite unclear what kind of AI (weak, strong, narrow, or AGI) and what relationship or role (human-supported AI, AI-supported human, or human-compatible AI) it is based on.

This could potentially have dire consequences, both on trust for the field of AIED and on the educational practise, in that it risks to contribute to hype and uncertainty in what to expect from AIED-systems in education.
Nevertheless, there are opportunities for AIED mentioned in all categories of the findings. If an AIED system is develop that is that strong that it can perform any task at the same level, or above, as a human teacher (strong AI or AGI), then why should it not take the lead or replace the human? Further, a well-designed narrow AIED system could, even though it does not possess intelligence in the human sense, support the teacher in many aspects of everyday teaching. Freeing the teacher from tedious and repetitive tasks, and making more time for the social and collaborative aspects of education.

The recommendation from the authors is that a more distinct terminology, inspired by recent research and discussion on AI, is adopted in the field of AIED to distinguish realistic applications from hopes, fears and science fiction. The authors would like to propose a terminology where weak and strong AI is combined with narrow AI and AGI to better describe the full range of artificial intelligence: weak-strong narrow AI and weak-strong AGI (Humble & Mozelius in press). Concerning which standpoint to take in the relationship between human and AI in the development and integration of AIED-systems, the AI-supported teacher is the most likely to be adopted today since it correlates with the technical level that we are at (narrow AI). However, it provides little preparation for the potential development of AGI and its possible consequences on education and therefore the authors would like to suggest the standpoint of teacher-compatible AI.

7. Future research

This study was carried out as a literature review on contemporary research in the field of AIED, and an interesting next step would be to further investigate the use of AIED systems in real-life educational practises in future research. An especially interesting focus would be on AIED systems potential effect on the learning outcomes for students, since that is the main aim for many of the systems discussed in this study.

References