# Evidence summary

# Artificial Intelligence in education

By Dr Carmel Kent,
Senior Research Fellow, EDUCATE

Artificial Intelligence (AI) is all around us - yet the media overwhelms us with a flux of contradictory narratives. Is AI a magic algorithm or a dangerous enemy? Is it causing a revolution or a disruption? Will it destroy, take over and suppress us - or will it augment, support and even free us? How do machines gain 'intelligence'? And most importantly, what will dictate the impact of AI on us humans? AI feels like a moving target. If there is one definitive fact about AI, it's that it will require us to learn throughout our lives.

The aim of this report is to summarise evidence about AI that is pertinent to education. Why education? Because to understand AI, we first need to understand human intelligence and human learning. We need to be able to identify the difference between AI and Human Intelligence (HI) if we are to reap the potential of AI for society. In addition, since our students and children will experience the greatest impact of AI - both from an employment perspective, but also from cultural and sociological perspectives - we need to evaluate how AI impacts education.

This is an overview of the main concepts that make up the image of AI today and explores the promise of AI in education. To do this, we must also discuss the challenges faced by entrepreneurs, designers, developers and policymakers in the field of AI in education. This will be the main aim of a future report. But let's begin by getting to know the enemy. Or, perhaps more appropriately, let's get acquainted with our new colleague.

## What (or who) is AI?

The best way to understand AI is to explore how it has come about. In the summer of 1956, John McCarthy (McCarthy & Hayes, 1969) initiated a two-month workshop, termed later as the Dartmouth College workshop, to "proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can, in principle, be so precisely described that a machine can be made to simulate it". McCarthy coined the term Artificial Intelligence (AI), and defined it as "the science and engineering of making intelligent machines that have the ability to achieve goals like humans do". This definition is instrumental to our understanding of AI today, through its contextualisation of machine intelligence as the ability to imitate the human processes of thinking and acting. Indeed, to understand the strengths and limitations of AI, it is imperative to also understand the similarities and differences between human cognitive systems and the machinery that typically attributes cognitive-like abilities to computers.

Machines have the advantage of being able continuously to scale their processing and storage capabilities, whereas humans have the advantage of being able to solve complex problems through a complicated (and sometimes hard to realise) interaction network of sensory cues, memory heuristics, emotions, experiences and cultural contexts. This complicated network of interactions has not yet been fully understood by any scientific discipline. As a result, machines cannot yet fully imitate the complex phenomenon of HI.

Another element of McCarthy's definition emphasises AI's interdisciplinarity. The magnificence of HI, and the way it has developed through learning, has caught the imagination of scientists from all disciplines. To really understand AI, and to evaluate its impact, it is important to recognise and explore the multidisciplinary scientific efforts made to get under the skin of human cognition, and the extraordinary efforts to formulate 'intelligence' using computational tools.

Looking at the interdisciplinary roots of AI, Russel and Norvig (2016) have extended McCarthy's definition to four forms of artificial achievement of human goals, as summarised in Figure 1, taken from their book.

| Thinking Humanly | Thinking Rationally |
|---|---|
| "The exciting new effort to make computers think … *machines with minds*, in the full and literal sense." (Haugeland, 1985)<br><br>"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning …" (Bellman, 1978) | "The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985)<br><br>"The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992) |
| **Acting Humanly** | **Acting Rationally** |
| "The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990)<br><br>"The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991) | "Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998)<br><br>"AI …is concerned with intelligent behavior in artifacts." (Nilsson, 1998) |

*Figure 1: Some definitions of AI, organised into four categories (Russel & Norvig, 2016)*

Our curiosity about intelligence stems from the Greek philosopher Aristotle (384–322 BC). Many philosophers wanted to understand what it is to 'think rationally', and it's a fascination that still impacts heavily on thinking about AI. Aristotle was one of the first people to try to formulate the rules of logic, depicting "right thinking", or irrefutable reasoning processes. This was the basis for the traditional deductive reasoning ("top-down logic"), in which conclusions are reached by applying general logical rules to observations. For example, when applying the logic rule 'all tomatoes are fruit' to an observed cherry tomato, we conclude that the cherry tomato is a fruit. Many of the first AI systems to appear - such as tutoring systems, or medical expert systems such as that for monitoring psychiatric treatment (Goethe & Bronzino, 1995) - were based on logical rules and deductive reasoning, because well-structured, formal rules are very easily coded into machine language. The advantage of this approach is that the logic of the system is clear, as well as being easy to follow and refute. The reasoning is transparent and understandable.

Sadly, these AI systems are very hard to maintain as the number of rules needed to implement a complex real-world problem can very quickly reach hundreds and thousands. In addition, translating informal knowledge into well-structured logical rules is a very challenging task – and such rules cannot deal with uncertainty. Gödel (1931) showed mathematically that deductive thinking is limited, and that there are always observations that cannot be obtained from generalised rules (Russel & Norvig, 2016). This questions the notion that everything can be computed by an algorithm in the form of a set of rules to be followed to reach a conclusion or solve a problem.

Deductive reasoning is not the only game in town, however. By contrast, inductive reasoning - as proposed by Hume (1739) in the form of what is now known as the principle of induction ("bottom-up logic") - describes universal rules that are acquired by generalising from exposure to repeated observations. For example, we might deduce that since the sun was shining every morning in the last week, it will also shine tomorrow morning.

The rise in the popularity of inductive reasoning is related to the philosophical movement known as empiricism, which is typically identified with John Locke's (1632–1704) quote: "Nothing is in the understanding, which was not first in the senses". This prepared the ground for the scientific practicalities we now sometimes take for granted: that cognitive models cannot theoretically be developed if no data and observations can support it.

As we will discuss in the next section, Machine Learning (ML) is most usually based on inductive reasoning, in the sense that ML models are developed on the basis of statistical patterns found in the observed data.

Returning to the limitations of deductive reasoning and the limitations of what can be computed by an algorithm, we find the motivation for Alan Turing's (1950) research to design the operational definition of intelligence as 'acting humanly' (Russel & Norvig, 2016). The famous Turing Test specified that a computerised system passes the test of 'acting humanly' if a human evaluator cannot tell whether a response was originated by a human being or by the system. Of course, this approach involves both the computing abilities of the machine and the expectations or perceptions of the human evaluator. This stresses again how AI is a moving target: if the evaluator's expectation of what an AI system can do changes, so does their educated guess in the Turing Test.

In their book, Russel and Norvig (2016) argue that to pass the Turing Test, a computer would need to possess:

- natural language processing (NLP) to enable it to communicate in human language
- knowledge representation (to store inputs and outputs)
- automated reasoning (to draw new conclusions deductively)
- machine learning (to adapt inductively to new circumstances)
- computer vision (to perceive objects)
- robotics (for movement)

With this definition, it is worth noting the theoretical limits of this algorithmic approach. Turing himself has shown that no machine could tell whether a given programme will return an answer on a given input, or run forever. Hence, AI systems are not as strong as humans in general-purpose settings, working more effectively on narrow and well-defined problems. In addition, a certain type of problem, called a NP-complete problem (Cook, 1971; Karp, 1972), is intractable; it cannot be solved computationally in 'reasonable time'.

An alternative approach to exploring the process of making decisions relates to humans' ability to 'behave rationally'. Do machines necessarily need to produce rational outcomes, given that humans do not always act rationally? This type of questioning brings us to Decision Theory (Raiffa, 1968), which provides a mathematical framework for analysing the decision process, designed to maximise the decision-maker's expected utility.

More recently, Game Theory (Binmore, 1992) has studied decision-making processes to maximise the utility of a decision-maker in an encounter with other decision-makers (see, for example, The Prisoner's Dilemma, from Tucker, 1950).

Both Decision Theory and Game Theory are based on the idea of a 'rational agent' - a prevalent concept in AI. A rational agent is one that acts so as to achieve their best-expected outcome, and thus will make the 'correct' logical inferences to maximise a known utility function. Kahneman &

Tversky (1982), two psychologists who won the Nobel prize in economics, showed that the assumption of humans acting and making decisions rationally is frequently incorrect. This raises interesting questions about the role of AI systems in imitating humans. Perhaps AI could usefully be employed to alert us about irrational decision making, instead of imitating our irrational behaviour?

## What does it mean to 'think humanly'?

Early research, led by John Watson (1878–1958) and described as 'behaviourism', proposes a systematic approach to understanding human learning. This approach gained some popularity in the first half of the twentieth century. Along with other methodological arguments, behaviourists argue that because the human cognitive process is 'unobservable', it cannot be studied - hence their focus is placed upon the analysis of behaviour. They believe this is the only observable evidence of human cognitive processes.

In contrast, cognitive psychology and neuroscience (the study of the nervous system, particularly the brain) gained much more traction in the second half of the twentieth century and led to much of our understanding of human cognition, and to our thinking about AI systems as 'thinking humanly'. Rashevsky (1936) was the first to apply mathematical models to the study of the nervous systems, showing that neurons (which are 'observable'), can lead to thought and action (Russel and Norvig, 2016). Most influential studies, such as Miller's (1956) Magic Number Seven and Chomsky's (1956) Three Models of Language, followed the cognitive psychology view of the brain as an information-processing device (Atkinson & Shiffrin, 1968), leading to their investigation of human cognition from a computational point of view.

To read more on the history of AI, the interested reader is referred to Russel & Norvig (2016), Smith et al. (2006), Kilani et al. (2018), Gonsalves, (2019) and Menzies (2003).

## From ELIZA to Go: the history of AI through five influential examples

Before we move on to discuss Machine Learning (ML) in detail, it is worth mentioning five famous AI systems from the past. These emphasise the evolution from deductive, rule-based AI to inductive, ML-based AI.

### Deductive-based AI: ELIZA, PARRY and RACTER (1960s – 1980s)

These are three of the early AI systems. ELIZA appeared in the 1960s, PARRY in the 1970s, and RACTER in the 1980s. These systems adopt a rule-based approach to natural language processing.

ELIZA was a text based conversational programme that presented itself as a Rogerian therapist (Weizenbaum, 1966). It was designed to show the superficiality of the communication between a human and a machine to which many people attributed human-like feelings. Rogerian therapies are abstractly based on conversations in which the therapist responds to the patient, reflecting back on their statements, and rephrasing them into questions. This basic conversational logic is well-suited to rule-based AI systems because they can use deductive logic to rephrase content created by the patient. If the patient's statement does not fit this rough logic (always in English), ELIZA can choose from a set of fixed phrases such as "Very interesting. Please go on." or "Can you elaborate on that?" (Güzeldere & Franchi, 1995). See for instance the example illustrated in Figure 2:

```
young woman: Men are all alike.
eliza: In what way?
young woman: They're always bugging us about something specific or other.
eliza: Can you think of a specific example?
young woman: Well, my boyfriend made me come here.
eliza: Your boyfriend made you come here?
young woman: He says I'm depressed much of the time.
eliza: I am sorry to hear you are depressed.
young woman: It's true. I'm unhappy.
eliza: Do you think coming here will help you not to be unhappy?
young woman: I need some help, that much seems certain.
eliza: What would it mean to you if you got some help?
young woman: Perhaps I could learn to get along with my mother.
eliza: Tell me more about your family.
young woman: My mother takes care of me.
eliza: Who else in your family takes care of you?
young woman: My father.
eliza: Your father.
young woman: You are like my father in some ways.
eliza: What resemblance do you see?
young woman: You are not very aggressive but I think you don't want me to notice that.
eliza: What makes you think I am not very aggressive?
young woman: You don't argue with me
```

*Figure 2: a conversation between ELIZA and a young patient (Güzeldere & Franchi, 1995)*

In 1972, just a few years later, Kenneth Colby created PARRY - a computer programme attempting to simulate a person with paranoid schizophrenia. A variation of the Turing Test was used with a group of psychiatrists who were shown transcripts of conversations between PARRY and human patients. These human psychiatrists were only able to correctly identify PARRY as a computer less than half of the time. PARRY was 'specialised' to elaborate on its 'beliefs, fears, and anxieties' in a question-answer mode. See in Figure 3, a conversation between ELIZA and PARRY (Güzeldere & Franchi, 1995):

```
parry: People get on my nerves sometimes.
eliza: I am not sure I understand you fully.
parry: You should pay more attention.
eliza: Suppose you should pay more attention?
parry: You're entitled to your opinion.
eliza: What makes you think I'm entitled to my opinion?
parry: It bothers me just to be around people.
eliza: Please go on.
[Parry continues with some paranoid ramblings about horseracing, bookies, and the Mafia
parry: I know the mob controls the big rackets.
eliza: Who are you really talking about?
parry: Have you ever really thought about the Mafia?
eliza: Suppose you ever really thought about the Mafia?
parry: Let's talk about something else.
eliza: I am not sure I understand you fully.
parry: You keep asking me the same things, why not change the topic?
eliza: In what way?
. . . . . . . . . .
parry: I am getting tired of answering the same questions over and over.
eliza: What resemblance do you see?
parry: Why do you need to know that?
eliza: What do you think?
...
```

*Figure 3: A conversation between ELIZA and PARRY (Güzeldere & Franchi, 1995)*

In the early 1980s, William Chamberlain and Thomas Etter, programmed RACTER (Chamberlain, 1984), the amusing 'artificially insane' raconteur. Below is a conversion between RACTER and ELIZA, and a short poem, written by RACTER (Güzeldere & Franchi, 1995). Unlike ELIZA, PARRY and RACTER gave the appearance of creating new content for the conversation, and thus display their own 'characteristics'.
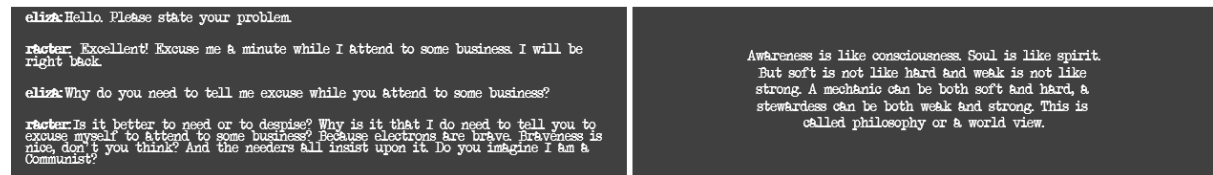


eliza: Hello. Please state your problem.

racter: Excellent! Excuse me a minute while I attend to some business. I will be right back.

eliza: Why do you need to tell me excuse while you attend to some business?

racter: Is it better to need or to despise? Why is it that I do need to tell you to excuse myself to attend to some business? Because electrons are brave. Braveness is nice, don't you think? And the needers all insist upon it. Do you imagine I am a Communist?

Awareness is like consciousness. Soul is like spirit. But soft is not like hard and weak is not like strong. A mechanic can be both soft and hard, a stewardess can be both weak and strong. This is called philosophy or a world view.

*Figure 4: Left: a conversation between ELIZA and RACTER; Right: a poem written by RACTER (Güzeldere & Franchi, 1995)*

**Game-playing AI: Introducing inductive AI (1990s until today)**

Early AI researchers were somewhat obsessed with chess. Unlike the 'conversational intelligence' that ELIZA, PARRY and RACTER aspired to, the intelligence associated with chess-playing is about strategy, planning, and cognition (Burgoyne et al., 2016). Thus, the development of an AI system that could play chess was seen as an intelligent goal to pursue.

Deep Blue is an AI system that defeated the world's chess champion in 1997. It used tree-search algorithms, which essentially traverse a hierarchical solutions' space (also called a game space) until finding an optimal solution. Tree-search algorithms are suited to a deductive logic approach, because the solutions' space is given in advance, and is specific to the game (or problem), rather than to the set of steps taken in real-time by the player. The huge solutions' space of Deep Blue was supported by IBM's massive-scaled hardware, which was able to support the inspection of 200 million board positions per second.

It is interesting to note that Google's AlphaGo, an AI system that defeated Lee Sedol, the world Go champion, in 2016, also used a tree-search algorithm. However, as this solutions' space was much larger, AlphaGo could not be efficiently supported solely by a deductive approach. Thus, Google's engineers used an ML neural network algorithm to reduce and optimise it beforehand, pre-calculating the most sensible moves for each possible board position. This optimisation, among others, enabled the tree-search algorithm to work efficiently in real-time and to defeat its human opponents.

Both these game-playing AI systems master a very specific task: to understand the current 'representation' of the game at every step, and quickly respond with the next best move. Every time they meet the same representation, they will produce exactly the same move. These systems cannot, by any means, transfer their mastery or adapt to new environments. They cannot play even a slightly different game or show any intellectual ability. Thus, according to McCarthy's definition, these two systems achieved their goal 'like humans do' and even outperformed us. However, having merely this single, non-transferrable skill cannot make them 'intelligent' in human intelligence terms.

## Machine Learning (ML): a sub-field of AI

Machine learning is a sub-field of AI, which is associated with machines' ability to learn inductively - that is, to "improve automatically through experience" as phrased by Tom Mitchell, one of the field's early contributors. Unlike Deep Blue, which is a purpose-built AI application programmed that reacts by following ready-crafted heuristics and rules, ML applications process sets of historical observations (data records) to infer new patterns or rules arising from the data itself. This approach challenges the concept of hardwiring the programming of specific behaviours. Whenever the data is changed, an ML algorithm 'learns', picking up the changed or modified patterns to present or predict a new result.

To better understand ML, consider this example of a medical decision-support system for hypertension that is built on programmed rules. If blood pressure after lunch, while sitting, is less than 140/90 and the patient has a family history of stroke, such a system is designed to recommend administering drug X. In this rule-based AI system, the currently treated patient's record is processed, and a recommendation is derived by deduction from the 'human knowledge' that has been programmed into the system.

As an alternative, an ML application would process all past patient records (with no previous coded human knowledge) and infer statistical patterns from them. For example: there may be a probability of 88% that patients who have a family history of stroke and have shown a positive response to drug Y would also respond positively to drug X, without any dependence on their age or current blood pressure level. The latter AI system is inductive: it has created a probabilistic rule out of the processed data. If the same system is trained (i.e., the process through which the system 'learns') on a set of patient records from a different hospital or country, it is likely to come up with (induce) a different set of probabilistic rules.

A further classic explanation of ML can be derived from exploring spam. By the early 2000s, as email usage gained momentum, the volume of spam was threatening to hurt email's efficacy. ML was the only approach that could learn and adapt quickly enough to the changes in spammers' tricks to manage the problem.

Another example can be seen in the way that [Google has transformed some of their technologies](#) (such as speech recognition and translation) using ML. ML frees AI from having to formalise and maintain coded human knowledge. This can, contextually, be either an advantage or a disadvantage – and, with ML, a dependency on coded human knowledge is replaced with a dependency on historical data. ML is generally very sensitive to the data it is trained on. If the data is inaccurate, irrelevant, insufficient or missing, an ML application will not be able to meaningfully induce rules or models from it.

In many senses, ML practitioners are, essentially, historians: they collect as many observations as possible about a historical event or behaviour and try to generalise reasoning from these observations. Like history, there is much political influence on the decisions that designers make about the observations or data they choose to collect, how to collect them, and how to interpret the induced models.

Unlike most historians, however, ML practitioners sometimes use these models to predict the future. Hence, it is important to understand that ML merely illustrates, perpetuates, amplifies and sometimes even simplifies past behaviour. When used as a decision support tool, it is important to

remember that ML will not, by itself, improve the ethics or biases already rooted within the data it was trained on (for example, the [Amazon discriminating recruitment AI](#)). Similarly, rule-based systems will perpetuate the mindset of the experts who built them, adding a layer of inference, but not a wind of change.

**Two types of the commonly used ML algorithms: supervised and unsupervised learning**

The term 'supervised learning' is used to describe ML algorithms that are trained on a data set that includes the outcome values. In our medical decision-support example above, the historical patient records include information about whether past patients had responded well to drug X. From this historically known outcome, the supervised learning ML algorithm can advise or predict the outcome for the new patient.

Another example of supervised machine learning is that of an image processing system trained on a set of images that has been annotated by humans to identify whether or not the image includes a car. The supervised ML algorithm will try to learn and predict whether a new, unannotated image includes a car. There are two main characteristics of supervised ML algorithms to consider here. Firstly, they are heavily dependent on human annotations of a large set of data. In cases where the outcome needs to be judged by experts to whom access is limited, this might be a problem. Secondly, the decision made by a supervised ML algorithm will be as biased as the human annotation it is drawn from.

'Unsupervised learning' is used when we do not have the values of outcomes at our disposal. That is to say, there is no 'human guidance' or supervision inherent to the algorithm. We would still, however, like to identify patterns hidden in the data. For example, we may want to find groups of similar ability students, in terms of their English and maths grades, in order to tutor them separately or to use different interventions with each group. Figure 5 shows four identified clusters that might be treated by teachers using different strategies.
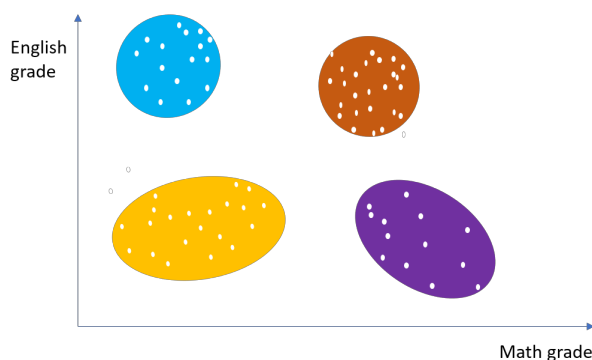


*Figure 5: Unsupervised ML resulting clusters*

A third, less common, type of ML is 'reinforcement learning'. Like supervised learning, this uses feedback to find and learn the 'correct behaviour'. Unlike supervised learning, however, reinforcement techniques do not use a given outcome as feedback. Instead, they use a set of rewards and punishments as signals for positive and negative patterns of behaviour.

To learn about ML, the interested reader is referred to Shalev-Shwartz & Ben-David (2014) and Witten, et al. (2016).

**AI in Education**

The research field of AI in education (AIEd) has existed for at least 30 years. AIEd "brings together AI … and the learning sciences … to promote the development of adaptive learning environments and other AIEd tools that are flexible, inclusive, personalised, engaging, and effective … AIEd is also a powerful tool to open up what is sometimes called the 'black box of learning,' giving us deeper, and more fine-grained understandings of how learning actually happens" (Luckin et al., 2016).

As an example, to develop an AIEd application providing individualised feedback to students, Luckin et al. (2016) argue that research from the learning sciences needs to be assimilated into three types of computational models: the pedagogical model (expressing teaching methods), the domain model (expressing the taught subject knowledge) and the learner model (expressing the personal cognitive, affective and behavioural attributes of learners). AI is only just starting to bring change to the educational ecosystem and as yet it has not necessitated that all educational stakeholders erngage with AI and its implications for education. However, education and educators need to prepare for the inevitable progress of AI into education. Luckin and Cukurova (under review) propose three main required actions to effectively connect AI and education, as summarised in Figure 6.



*Figure 6: Luckin and Cukurova's intelligent approach to AI in education and training*

We will discuss the first two actions in the next section, and the third action in a future publication.

**Human or machine learning?**

Skinner (1938), one of the most influential behaviourist psychologists, developed a learning method based on the notion that people learn when they adopt an association between a particular behaviour and its consequence (either a reward or a punishment). When the learner associates certain behaviour with a reward, for example, they are likely to repeat it.

These ideas share some similarity to the ML's supervised and reinforcement methods as we know them today, in which a statistical model is 'taught' by associating an observation with a consequence (reward, punishment, or simply an already known outcome).

The reasons that behaviourist methods are less acceptable today are similar to the reasons that human learning is so different from that of a machine: there is human cognition sitting in between, which is more complicated than simply responding to rewards. The confusion between human and artificial intelligence is not just seen in the analogy drawn between machine and human cognition by Russel and Norvig (2016) as illustrated in Figure 1 above, but also in the Turing Test, and the information-processing analogy (Atkinson & Shiffrin, 1968). All these examples similarly ignite the imagination of scientists and engineers to explain both human and computational systems while prompting utopian misconceptions, some of which cause anxiety about AI.

Armed with the basic understanding of AI and ML, the rest of this sub-section tries to pinpoint some of the main differences between human cognition and machine cognition that we currently understand. We take a risk when writing anything down here that readers will view the situation as fixed, when this should really be a living document, because AI's abilities are still being continually developed. It is therefore beneficial to frame our discussions of each system through its strengths and weaknesses, so that we can try to shape the future of this complicated, evolving relationship.

**No free lunches**

Current AI systems, whether rule-based or ML, cannot address problems they were not trained for or designed for. We have well-trained Go playing machines, self-driving cars and automatic image recognition system for cats – but we still lack a 'well-rounded' machine learner.

Wolpert & Macready's (1997) 'no free lunch' theorem explains why there is no general-purpose artificial cognition, explaining: "if an algorithm does particularly well, on average, for one class of problems then it must do worse, on average, over the remaining problems."

It is not unthinkable that AI systems that can cope with a range of different models to suit a range of different problems will emerge in the future. These would most probably be dependent on access to extremely rich and multidimensional datasets, which should be maintained. However, computing and implementing the interactivity between these models is still hard to imagine.

There are caveats to any form of AI being used in isolation. Rule-based or algorithmic-based AI systems (as opposed to ML) - such as Deep Blue and the earlier expert systems – can be optimised for routine tasks (Levy & Murnane, 2012). The deductive approach, in which intelligence is perceived as a form of algorithmic computation (i.e., a programmed set of computations and rules) are optimised for repetitive, well-structured and well-defined tasks. However, when uncertainty, complexity and change are introduced, and some form of creativity and adaptation is needed, these AI systems fall short. This caveat led to the gain in popularity of the inductive approaches to AI used by ML.

Having said this, there are caveats to pure ML too. ML systems cope well with change, since they are inductive: once a set of observations changes, the derived conclusion or prediction changes as well. However, ML is limited to the kind of conclusions drawn from existing observations. Unlike the human brain, machines cannot solve problems to which they have not previously been introduced. As Russel (1997) emphasises: "imagine a chicken that gets fed by the farmer every day and so, quite understandably, imagines that this will always be the case… until the farmer wrings its neck! The chicken never expected that to happen; how could it? – given it had no experience of such an event and the uniformity of its previous experience had been so great as to lead it to assume the pattern it had always observed (chicken gets fed every day) was universally true. But

the chicken was wrong". In other words, AI inductive systems will not consider any choice of action to which historical evidence was not introduced. Popper (1968/2002) argued that while observations and experimentation play an important role in knowledge creation, the emphasis of science should be on finding evidence for falsifying the induced conclusions, and not assume that these induced conclusions are correct just because they have not (yet) proved otherwise. ML isn't currently designed to do this.

Humans are inherently 'designed' to do both deductive and inductive learning. As we collect observations through our senses and process them to fit into our long-term memory schemas, we are inducing. On the other hand, as we use heuristics and our long-term schemes and scripts, we are deducing predictions and possible explanations for our observations (e.g., Atkinson & Shiffrin, 1968).
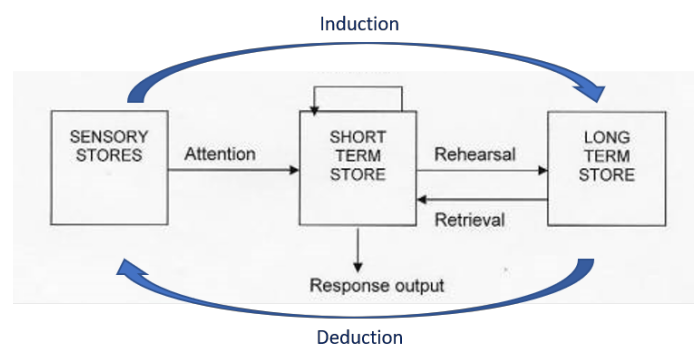


*Figure 7: Human induction and deduction, adapted from Atkinson & Shiffrin (1968)*

## Learning by imitation is not enough

"The quest for 'artificial flight' succeeded when the Wright brothers and others stopped imitating birds and started using wind tunnels and learning about aerodynamics" (Russel & Norvig, 2016). AI systems are good at picking up patterns, repeating and generalising them - but they lack creativity and cannot transfer skills. Luckin (2018a) named seven elements to human intelligence, that still do not have any complete analogue in artificial cognition:

1. **Multi and interdisciplinary academic intelligence**, described as knowledge and understanding about the world.
2. **Social intelligence**. Luckin explains: "social interaction is the basis of individual thought and communal intelligence. AI cannot achieve human-level social interaction. There is also a meta aspect to social intelligence (see also meta-subjective Intelligence) through which we can develop an awareness of our own social interactions and hone our ability to regulate them."
3. **Meta-knowing intelligence**. The understanding of "what knowledge is, what it means to know something, what good evidence is and how to make judgements based on that evidence and our context."
4. **Meta-cognitive intelligence**. The ability to "interpret our own ongoing mental activity: interpretations that need to be grounded in good evidence about our contextualised interactions in the world."
5. **Meta-subjective intelligence**. Encompasses "both our emotional and our motivational self-knowledge and regulatory skills; our ability to recognise our emotions and the emotions of

others; to regulate our emotions and behaviours with respect to other people and with respect to taking part in a particular activity."

6. **Meta-contextual intelligence**. Described as "our understanding of the way in which our physical embodiment interacts with our environment, its resources, and other people. This includes physical intelligence; our intellectual bridge to our instinctive mental processes. This helps us recognise when we are biased and when we are succumbing to post-hoc rationalisation."

and most importantly, connecting all the above six elements:

7. **Perceived self-efficacy**. Requiring "an accurate, evidence-based judgement about ourselves: our knowledge and understanding; our emotions and motivations; and our personal context. We need to know our ability to succeed in a specific situation and to accomplish tasks both alone and with others."

Luckin suggests that the human ability to reflect about our learning, and to understand and process contextual and subjective knowledge through experience, is the core difference between human cognition and machine cognitions.

While AI systems are unable to reflect upon themselves and be self-aware, AI systems are undoubtedly superior at capacious storage and processing speed. The memory device called the Internet and the massive acceleration in computing power has equipped AI with an unfair advantage. AI is at its best when recalling, searching and recognising.

**Human heuristics and cognitive biases**

To deal with our limited processing capacity, memory loss, and memory decay while still making sense of the world, humans use heuristics, schemas and scripts. Heuristics are mental shortcuts people often use to make decisions, usually focusing on just a few aspects of a situation (for example, rule of thumb, educated guesses and stereotypes). Schemas and scripts are mental structures for preconceived ideas and known processes that people use to understand the world. For example, a script informing you of what you should expect when entering a restaurant will probably involve being seated, being given a menu and expecting a waiter to arrive. Tversky & Kahneman (1974) showed how such mental shortcuts often lead us to mistaken probabilistic conclusions.

Fortunately, machines do not need to use such shortcuts. Even computational methods that are used to reduce the number of considered dimensions and aspects (such as feature selection) are based on statistics. Thus, machines could help us to identify biases and make better-informed decisions.

AIEd focuses on ways in which human deficiencies could be complemented by machine abilities and vice versa, rather than seeing human cognition and machine cognition as supplementary and in conflict. For example, teaching machines could operate by breaking human-understood tasks into simpler machine-understood tasks (Azaria et al., 2016).

**AI in education**

Skinner was not only one of the forefathers of behaviourism, he was also one of the first edtech (educational technology) entrepreneurs. He identified the problem of parents' workload when his second child was born, and developed a new technology designed to solve it: the "air crib" (Skinner, 1945). In his article 'Baby in the box', he explains: "I felt that it was time to apply a little labour-saving invention and design to the problems of the nursery. We began by going over the disheartening schedule of the young mother, step by step... Then the 'gadgeteering' began."

With Skinner's design, the baby spends all its time in the 'air crib', except for "about one and one-half hours each day to feed, change, and otherwise care for the baby" (Skinner, 1945). For the rest of the time, the crib supplies all the baby needs: a controlled temperature and a germ-safe climate.

Like many other edtech solutions that are built primarily with the technology in mind, the crib was controversial and was not adopted. Fortunately, the AIEd research community generally argues for a 'pedagogy first' approach, in which edtech innovations undergo a thorough exploration of the educational problems and gaps for which the technology will be tailored (Rosé et al., 2018).

In this report, we focus on three broad educational areas, and point to some of the work that has been done using AI to address each one. The next three sub-sections will therefore focus on how AI solutions are augmenting learning, teaching, and assessment.

**Augmenting learning**

Most formal classrooms are still based on an industrial-age factory model, almost intact. Most students - whether in primary, secondary or higher education - still learn in (large) heterogenous groups, generally organised by age stratification, over a certain pattern of time and space, progressing through terms and school years towards meeting standardised assessment criteria and qualifications (Bates, 2015). This structured organisation poses many challenges to the individual learner, to which AI - using automation (mostly via the deductive approach) and adaptivity (mostly via the inductive approach) - is being used.

**Personalised and adaptive learning**

In most formal education settings, the majority of individuals are taught as part of a large group of learners, all holding a highly diverse set of skills, abilities, contexts and interests - all facing a single teacher. Unlike teachers, AI systems can scale very easily and quickly, and can be used to facilitate a one-on-one interface with a learner, taking into account a large number of sensory inputs in real-time, and calculating on-the-spot recommendations for the most suited content, pace or instruction method for that specific learner at that specific time. These recommendations can be given either within or outside the standard curriculum (which is usually programmed 'top-down' into the system).

Personalised learning relates to the tailoring of learning resources or methods to fit with the specific cognitive, affective and behavioural needs of each individual learner - or to simply provide the right feedback. For example, personalised systems can take advantage of AI's computational abilities to consider many inputs about the learner within a single statistical model, resulting in a single recommendation about the next step.

Adaptive learning (which is often implemented alongside personalised learning) is about an AI system's ability to adapt in real-time to the dynamically changing needs of the learner. Adaptability can be gained, for example, by harnessing ML's ability to re-craft a statistical model from newly introduced data.

Intelligent Tutoring Systems (ITS) use AI techniques to simulate one-to-one human tutoring, usually using personalised and adaptive learning. Examples include Kidaptive (https://kidaptive.com/), which collects a diverse set of measures and uses AI to adapt content and feedback to the learners; ALEKS (www.aleks.com/), which uses a diagnostic approach throughout the learning journey to provide each learner with recommended topics; IBM and Pearson's cognitive tutor (www.ibm.com/watson/education/pearson); and CENTURY Tech (www.century.tech), which collects behavioural and performance data to recommend the next step, while also providing tracking analytics for teachers and auto-marking to give instant feedback to students (Luckin et al., 2016).

The academic literature on the efficacy of adaptive and personalised learning is not unequivocal. It is important to remember that it is not trivial to evaluate the efficacy of technological solutions in the context of a system that evaluates outcomes most often using paper-based exams, using standardised criteria. However, Basitere & Ivala (2017) for example, did find that the use of a personalised adaptive system positively impacted on students' performance in paper-based tests in physics. On the other hand, Reich and Ito (2017) argue that a significant effect of using adaptive systems as compared to traditional instruction is not often found and suggest that what AI assisted systems excel at is evaluating 'computational skills', which mostly resemble the AI's way of processing.

Ease of communication AI-based translation technologies (such as those provided by Microsoft or Google) empower learners across the globe to consume high-quality content (for example, by producing real-time subtitles to lectures) or language and culture learning driven by virtual role-playing. For example, Alelo (www.alelo.com) uses AI in the form of experiential learning through the Tactical Language and Culture Training System (TLCTS), which uses social simulation models for language learning (Johnson & Valente, 2009). AI could also be used to help students in writing. For example, the Academic Writing Analytics (AWA) system (Gibson et al., 2017) is a web-based system, automating reflective writing analytics to provide formative feedback on students' writing.

**From teaching machines to tutors' systems to augmenting teachers**

Almost side-by-side with the term AI, the counterpart term Intelligence Augmentation (IA) (van Emden, 1991) has developed. While AI traditionally pushed towards autonomous systems that would eventually replace human cognitive functions, IA aims to use similar techniques to support humans by complementing cognitive functions, rather than replacing them. This change of tone emphasises the evolution of the relationship between AI (or IA) and the teacher's roles.

At 1924, Ohio State University professor Sidney Pressey - and later Skinner (1961) - invented a prototypical device intended to 'teach' students. Pressey's device, the 'automatic intelligence testing machine', was 'testing' the students by presenting them with multiple choice questions and letting them click on the right answer, or 'teaching' them by not revealing the next question before they got the previous one correct (Watters, 2015). Skinner's device, called the 'teaching machine', differed in that it exposed the students to new concepts: "A student is 'taught' in the sense that he

is induced to engage in new forms of behavior and in specific forms upon specific occasions" (Skinner, 1961).

Skinner stressed that "Education must become more efficient… In any other field a demand for increased production would have led at once to the invention of labor-saving capital equipment."

This notion of automating teaching has evolved, appearing in many early intelligent tutoring systems, such as SCHOLAR (Collins et al, 1975). This is a development that was criticised for focusing on technology rather than pedagogy, and for focusing on a very narrow set of teaching methods (Rosenberg, 1987).

Today, education tends to favour IA and systems that assist teachers over systems that actually 'teach' for them. IBM's Watson Teacher Advisor (https://teacheradvisor.org/landing), for example, aims to reduce tutors' workload. Writing about the AI Teaching Assistant, Colin, Luckin and Holmes (2017c) note: "Through working with Colin, [the teacher] has become somewhat of a metaphorical judo master, harnessing the data and analytical power of AI to tailor a new kind of education to each of her students. Her role at the helm of the classroom, however, is fundamentally unchanged… From time to time, when Colin recognises that a group is off topic, he intervenes with an alternative suggestion to stimulate new discussion, via individual students' tablets, or he links students to other conversations that are taking place elsewhere in the classroom. Meanwhile, [the teacher] is free to wander around the room and observe, giving personalised guidance and feedback, and joining in with students' conversations. By now, she is an experienced and skilled problem-solving practitioner, attuned to recognising when her human help and social skills are particularly needed."

**Augmenting assessment**

Learning assessment is a crucial process, tightly coupled to learning itself. It should be designed to ensure that learners are making progress towards acquiring the knowledge and skills targeted by the learning system. The Western world's prevailing assessment methods are critiqued by learning scientists and educators, mainly aiming at what is assessed and how.

*What is assessed*

Watson (2017) reflects: "What counts is whether you can regurgitate a series of facts and apply them in a logical manner that is consistent with the views of the examiner or exam board. At its most basic level, it's a memory test. At a more sophisticated level (and in later years of education) it's a test of understanding - but rarely do the tests assess anything other than the idea that every problem has a right answer."

Our examination system excels at assessing numeracy and factual knowledge, but falls short in assessing other skills, such as creative problem-solving, empathy, and collaboration (Luckin, 2017b). In other words, the assessment system perpetuates intelligence that resembles that of a machine instead of celebrating and encouraging diverse and rich human intelligence.

The examination system is not beyond politics as well. It rewards certain types of skills, certain subjects even, and therefore encourages a certain type of student. Luckin (2017b) argues that instead of rewarding humans for displaying skills we can easily automate, we should encourage the non-cognitive skills that differentiate us from machines. To the argument that these skills are hard

to assess, she answers that this is exactly where AI can help (Luckin, 2017b). For example, Competency.AI (https://competency.ai) uses AI to monitor the progress of medical students' competencies, based on their curriculum.

### *How we assess*

Overall, the 'single-point-in-time' exam method for assessment has proven less than optimal. Students are required to artificially stop their learning process and be assessed under pressure - most often without any formative and immediate feedback. Teachers are required to compromise a rigorous and dynamic evaluation of their students' knowledge and understanding, in which they are very rarely able to identify individuals' needs (Luckin, 2017a). Luckily, advances in AI data collection and modelling techniques can significantly contribute to providing "a fairer, richer assessment system that would evaluate students across a longer period of time and from an evidence-based, value-added perspective" (Luckin, 2017a). AI is already substantially used for plagiarism detection, and even for the automatic identification of writing styles (e.g., Emma, https://emmaidentity.com). Automatic essay-scoring uses AI to provide scaled and real-time feedback to students (Santos et al., 2012) and to classify text for any other purpose (Bayesian Essay Test Scoring System, http://echo.edres.org:8080/betsy). AI is also used to identify misconceptions by using Natural Language Processing (NLP) (Nye et al., 2014), and to provide students with timely personalised feedback (OnTask, www.ontasklearning.org, Pardo et al., 2018).

To read more about AIEd, the interested reader is referred to Luckin (2018a) and Luckin (2018b).

## Conclusions

Our report opened with the 1956 Dartmouth workshop - whose aim, as documented in the project proposal (see Figure 8) was to "write a calculator programme that can solve intellectual problems as well as or better than a human being."
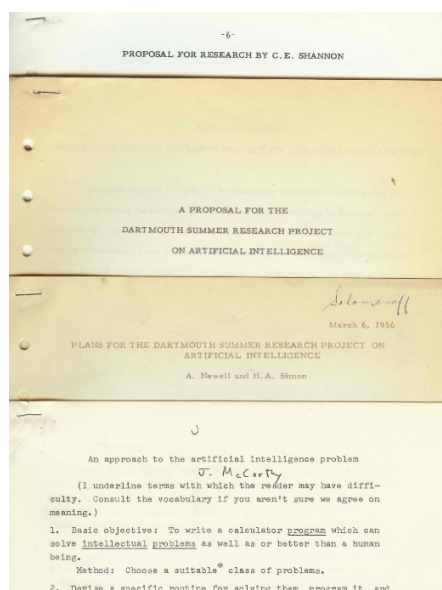


*Figure 8: The Dartmouth's proposal, as document by Ray Solomonoff, one of the participants of the workshop, at http://raysolomonoff.com/dartmouth*

Although this document is a few decades old, its basic axiom positions AI in comparison to humans, and this still resonates within the current extensive discussion around autonomous AI, within the design of AI technologies, in education, and in other areas.

Another definition states that AI is about 'computer systems able to perform tasks normally requiring human intelligence' (Oxford Reference, 2018). In the 1970s, when the first pocket calculators were shown to the world, they were considered a form of AI, because mathematical calculations were thought to require human intelligence. This prompts one to ask: which other daily routines will not be restricted to our human ability in the foreseeable future? Will AI be able to teach, learn and process complex decisions?

Like McCarthy's definition and Dartmouth's objective, the Turing's Test also positions machine cognition in comparison to that of a human: a machine exhibits 'intelligent' behaviour if it is indistinguishable (according to human criteria) from that of a human. Following the logic of Turing's Test, if a human process of problem-solving is not clear, then an AI system cannot be bullet-proof tested or validated. AI is still a science in the making.

This report reviews the concept of 'cognition comparison' from the point of view of the learning sciences. We suggest that human and machine means of processing information are rooted and developed very differently. Therefore, if we intend to refer to the intelligence of computational systems from the point of view of human cognition, we should reconsider the aspiration for machines to be 'like humans' and instead consider developing systems 'for humans', and according to humans' moral and educational values and needs.

The notion that machines augment rather than replace humans was strongly voiced by Doug Engelbart in the mid-20th century. His ideas of augmenting human intellect were still considered radical decades later (Engelbart, 1962, 2001), but were adopted by some thought leaders such as Rheingold (1985, 2000). In his discussion of 'mind-amplifying' technologies, Rheingold noted: "You can't really guess where mind-amplifying technology is going unless you understand where it came from."

We echo this statement, stressing that we must understand human learning to understand how machine learning can amplify it. 'Mind amplifying' in AIEd terms means an effective and egalitarian assessment system, with teachers empowered by their Colin-like assistants, to work with empowered, self-regulating, engaged students.

## References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes1. In Psychology of learning and motivation, 2, 89-195. Academic Press.

Azaria, A., Krishnamurthy, J., & Mitchell, T. M. (2016). Instructable Intelligent Personal Agent. In AAAI, 2681-2689.

Bates, T. (2015). Teaching in a digital age: Guidelines for designing teaching and learning for a digital age. Tony Bates Associates.

Basitere, M., & Ivala, E. (2017). Evaluation of an adaptive learning technology in a first-year extended curriculum programme physics course. South African Computer Journal, 29(3), 1-15.

Binmore, K. (1992), Fun and Games: A Text on Game Theory. D. C. Heath and Company: Lexington, MA.

Burgoyne, A. P., Sala, g., Gobet, f., Macnamara, B. N., Campitelli, G. & Hambrick, D.Z. (2016), The relationship between cognitive ability and chess skill: A comprehensive meta-analysis. Intelligence.

Chamberlain, W. (1984). The Policeman's Beard is Half Constructed: Computer Prose and Poetry. Warner Books.

Chomsky, N. (1956). Three models for the description of language. IRE Transactions on information theory, 2(3), 113-124.

Collins, A., Warnock, E. H., & Passafiume, J. J. (1975). Analysis and synthesis of tutorial dialogues. In G. H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (Vol. 8, pp. 49-87). New York: Academic Press.

Cook, S. A. (1971). The complexity of theorem- proving procedures. In STOC-71, 151–158

Engelbart, D. C. (2001). Augmenting human intellect: a conceptual framework (1962). PACKER, Randall and JORDAN, Ken. Multimedia. From Wagner to Virtual Reality. New York: WW Norton & Company, 64-90.

Gibson, A., Aitken, A., Sándor, Á., Buckingham Shum, S., Tsingos-Lucas, C., & Knight, S. (2017). Reflective writing analytics for actionable feedback.

Gödel, K. (1931). Uber¨ formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme I. Monatshefte für Mathematik und Physik, 38, 173–198.

Goethe, J. W., & Bronzino, J. D. (1995). An expert system for monitoring psychiatric treatment. IEEE Engineering in Medicine and Biology, November/December, 776–780.

Gonsalves, T. (2019). The Summers and Winters of Artificial Intelligence. In Advanced Methodologies and Technologies in Artificial Intelligence, Computer Simulation, and Human-Computer Interaction, 168-179. IGI Global.

Güzeldere, G., & Franchi, S. (1995). Dialogues with colorful "personalities" of early AI. Stanford Humanities Review, 4(2), 161-169.

Johnson, W. L., Valente, A. (2009). "Tactical Language and Culture Training Systems: Using AI to Teach Foreign Languages and Cultures". AI Magazine. 30(2), 72.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. Scientific American, 246(1), 160-173.

Karp, R. M. (1972). Reducibility among combinatorial problems. In Miller, R. E. and Thatcher, J. W. (Eds.), Complexity of Computer Computations, 85–103. Plenum.

Kilani, A., Hamida, A. B., & Hamam, H. (2018). Artificial Intelligence Review. In Encyclopedia of Information Science and Technology, Fourth Edition, 106-119. IGI Global.

Levy, F., & Murnane, R. J. (2012). The new division of labor: How computers are creating the next job market. Princeton University Press.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education.

Luckin, R. (2017a). Towards artificial intelligence-based assessment systems. Nature Human Behaviour, 1(3), 0028. https://doi.org/10.1038/s41562-016-0028

Luckin, R. (2017b). The Implications of Artificial Intelligence for Teachers and Schooling. In L. Loble, T. Creenaune, & J. Hayes (Eds.), Future frontiers : education for an AI world,109. Melbourne University Press & New South Wales Department of Education.109-125

Luckin, R. & Holmes, W. (2017c). A.I. Is the New T.A. in the Classroom, Available at: https://howwegettonext.com/a-i-is-the-new-t-a-in-the-classroom-dedbe5b99e9e

Luckin, R. (2018a). Machine Learning and Human Intelligence: The future of education for the 21st century. UCL IOE Press.

Luckin, R. (2018b). Enhancing Learning and Teaching with Technology: What the Research Says. UCL IOE Press. UCL Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL.

McCarthy, J., & Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. Michie D. Machine Intelligence., 463. https://doi.org/10.1016/B978-0-934613-03-3.50033-7

Menzies, T. (2003). 21st-century AI: proud, not smug. IEEE Intelligent Systems, 18(3), 18-24.

Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. The psychological review, 63, 81-97.

Mitchell, T. M. (2006). The discipline of machine learning (Vol. 9). Pittsburgh, PA: Carnegie Mellon University, School of Computer Science, Machine Learning Department.

Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. International Journal of Artificial Intelligence in Education, 24(4), 427-469.

Oxford Reference. (2018). Artificial Intelligence. Retrieved from http://www.oxfordreference.com/view/10.1093/oi/authority.20110803095426960

Pardo, A., Lim, L. & Yi-Shan, T., (2018), Connecting Data with Student Feedback in a course, In ALASI18 Australian Learning Analytics Symposium, Monash University, Melbourne, Australia.

Popper, K. R. (1968/2002). The logic of scientific discovery. Routledge.

Pressey, S. L. (1926). School and Society.23, 586.

Raiffa, H. (1968) Decision Analysis: Introductory Lectures on Choices under Uncertainty. Addison Wesley, Reading, MA.

Rashevsky, N. (1936). Physico-mathematical as- pects of excitation and conduction in nerves. In Cold Springs Harbor Symposia on Quantitative Biology. IV: Excitation Phenomena, 90–97

Reich, J and Ito, M. (2017). From Good Intentions to Real Outcomes: Equity by Design in Learning Technologies. California: Digital Media and Learning Research Hub.

Rheingold, Howard (2000) [1985]. Tools for thought: the history and future of mind-expanding technology (Reprint ed.). Cambridge, MA: MIT Press. ISBN 0262681153. OCLC 43076809

Rosé, C. P., Martínez-Maldonado, R., Hoppe, H. U., Luckin, R., Porayska-pomsta, M. M. K., Mclaren, B., … Goebel, R. (Eds.). (2018). Artificial Intelligence in Education II. In 19th International Conference, AIED 2018 London, UK, June 27–30, 2018 Proceedings, Part II (p. 580). https://doi.org/10.1007/978-3-319-61425-0

Rosenberg, R. (1987). A critical analysis of research on intelligent tutoring systems. Educational Technology, 27(11), 7-13.

Russell, B. (1997). Religion and science (No. 165). Oxford University Press, USA.

Russel, S., & Norvig, P. (2016). Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited. https://doi.org/10.1017/S0269888900007724

Santos, V. D., Verspoor, M., & Nerbonne, J. (2012). Identifying important factors in essay grading using machine learning. International Experiences in Language Testing and Assessment—Selected Papers in Memory of Pavlos Pavlou, 295-309.

Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge university press.

Skinner, B. F. (1938). The Behavior of organisms: An experimental analysis. New York: Appleton-Century.

Skinner, B. F. (1945). Baby in a box. Ladies Home Journal, 62(10), 30-31.

Skinner, B.F. (1961). Teaching machines, Scientific American, 205, 90–112.

Smith, C., McGuire, B., Huang, T., & Yang, G. (2006). The history of artificial intelligence. University of Washington, 27.

Tucker, A. W.  (1950). A two-person dilemmma, mimeo, Stanford University.

Turing, A. (1950). Computing machinery and intel- ligence. Mind, 59, 433–460.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. science, 185(4157), 1124-1131.

van Emden, M. H. (1991). Mental Ergonomics as Basis for New-Generation Computer Systems. University of Victoria, Department of Computer Science

Watson, R. (2017). On Education in the 21st Century. In L. Loble, T. Creenaune, & J. Hayes (Eds.), Future frontiers : education for an AI world,109. Melbourne University Press & New South Wales Department of Education.109-125

Watters, A. (2015). The Automatic Teacher, Available at http://hackeducation.com/2015/02/04/the-automatic-teacher

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine, Communications of the ACM, 9(1), 36-45.

Witten, I.H. Frank, E., Hall, M.A. and Pal, C.J., 2016. Data Mining, Practical Machine Learning Tools and Techniques, Morgan Kaufmann Series in Data Management Systems, 4th Edition.

Wolpert, D., & Macready, W. (1997). No free lunch theorems for optimization. IEEE transactions on evolutionary computation, 1(1), 67–82. doi:10.1109/4235.585893