Human Arenas
https://doi.org/10.1007/s42087-022-00304-8

ARENA OF TECHNOLOGIES

Scaffolding Human Champions: AI as a More Competent Other

Henrik Skaug Sætra[1]

Received: 28 February 2022 / Revised: 9 July 2022 / Accepted: 27 July 2022

Abstract
Artificial intelligence (AI) has surpassed humans in a number of specialised intellectual activities—chess and Go being two of many examples. Amongst the many potential consequences of such a development, I focus on how we can utilise cutting edge AI to promote human learning. The purpose of this article is to explore how a specialised AI can be utilised in a manner that promotes human growth by acting as a tutor to our champions. A framework for using AI as a tutor of human champions based on Vygotsky's theory of human learning is here presented. It is based on a philosophical analysis of AI capabilities, key aspects of Vygotsky's theory of human learning, and existing research on intelligent tutoring systems. The main method employed is the theoretical development of a generalised framework for AI powered expert learning systems, using chess and Go as examples. In addition to this, data from public interviews with top professionals in the games of chess and Go are used to examine the feasibility and realism of using AI in such a manner. Basing the analysis on Vygotsky's socio-cultural theory of development, I explain how AI operates in the zone of proximal development of our champions and how even non-educational AI systems can perform certain scaffolding functions. I then argue that AI combined with basic modules from intelligent tutoring systems could perform even more scaffolding functions, but that the most interesting constellation right now is scaffolding by a group consisting of AI in combination with human peers and instructors.

Keywords Artificial intelligence · Zone of proximal development · Scaffolding · Tutoring · Chess

Introduction

Human beings are no longer the masters of all intellectual pursuits. While artificial intelligence (AI) is incapable of challenging the breadth and adaptability of human intelligence, it has thoroughly defeated us at a broad array of specific tasks, such as playing *chess* and

✉ Henrik Skaug Sætra
Henrik.satra@hiof.no

1 Faculty of Computer Science, Engineering and Economics, Østfold University College, Remmen, 1757 Halden, Norway

Go (Fulbright, 2020). There are various perspectives on the consequences of such a development. We might argue that machine excellence is irrelevant for human activity, or we could see the consequences as either positive or negative for us. One consequence could be intellectual atrophy, but it could also be a diversification of the activities to which humans direct their intellectual vigour (Carr, 2020; Danaher, 2019; Strogatz, 2018).

By examining individuals in a particular position—the human *champions*—I show that machine excellence is *not* irrelevant. While human champions would previously have been the best entities in existence, they are now merely the best of their kind—*human*kind. A perennial challenge for human champions is to find competent tutors, as no humans who excel them exist. This, however, is where AI surpassing human abilities potentially radically changes things. I here emphasise the new possibilities that emerge between the best of humanity and the machines that surpass them in specific activities. The idea of combining human and computer powers in order to achieve progress is not new, and many have emphasised the possibility of learning from AI (Heaven, 2019; Kasparov, 2019; Strogatz, 2018). However, the *process* by which such learning occurs remains underdeveloped.

I propose that using the theory of Lev Vygotsky (1896–1934) helps us understand *how* AI can promote human learning. When a human champion is surpassed, a phenomenon unique in human history occurs, as machines exist in our champion's zone of proximal development (ZPD). Drawing on Vygotsky's ideas of scaffolding and learning from more competent others in this zone, I examine the potential for AI to accelerate and aid human intellectual development. I do so by using the examples of two high-profile games: chess and Go. In the worlds of both of these games, AI has become a phenomenon of huge importance, as it has thoroughly conquered these games and the dethroned humans.

The general framework of an AI powered learning system here provided serves as a theoretical proof of how AI can be developed to encourage the use of AI in systematic scaffolding where no other humans are available as the more competent others. When cutting-edge AI is combined with traditional theories of learning and conventional educational technologies, it may play the role of the more competent partner, or *other*, of our champions. I will show how some functions of scaffolding can be performed already, while others require the further development of explainable AI. Additional steps may then be taken by combining cutting-edge AI with elements from intelligent tutoring systems (ITS). Today, however, the bleeding edge of expert knowledge exists in teams where the best of AI is combined with human experts—*jointly* scaffolding our human champions.

Firstly, the theoretical framework is presented, which results in the presentation of the AI-based expert scaffolding system. Secondly, the framework is examined through the examples of interviews and quotes from top professionals and experts in the games of chess and Go, in order to examine the validity of the claims made about AI's potential. Thirdly, the current and future potential of AI as tutors of human champions is discussed on the basis of the preceding sections.

Theoretical Framework

This is a theoretical exploration of how advances in AI can be combined with theories of human learning and the use of educational technology in order to advance human knowledge. Firstly, I present a background of work on the use of computers and AI in educational settings. Secondly, Vygotsky's theory and the core concepts I rely on is presented. This theory serves as a key element in the education system proposed, and it also serves to show

how such an education system functions in the novel situation where human champions take the role of learners. Lastly, the outline of an AI-based expert scaffolding system is proposed, and this serves as the theoretical foundation of the article.

Computers in Education

Computers and various forms of AI have been applied in educational settings for a long time. Since the 1950s, computer-assisted learning/instruction applications have developed from basic linear programs to the adaptive and advanced educational system used today (Nwana, 1990). Since the 1970s, AI in education (AIED) has emphasised how AI can be used to analyse learners and tailor instructions based on the resulting data (Humble & Mozelius, 2019; Nwana, 1990). This is one of the aspects discussed in the framework here proposed, but not the only one, as AI can also take on the role of the expert and the producer of learning materials.

Brusilovsky (1999) describes how AIED systems, now commonplace, were based on earlier technologies such as *intelligent tutoring systems* (ITS) and *adaptive hypermedia systems*. As my main emphasis is on the use of AI in tutoring functions, I will use the term ITS for the foundational framework here developed. Such systems contain knowledge about the subject matter to be taught, the learner, and the path to achieve learning (Nwana, 1990). While Erümit and Çetin (2020) refer to *adaptive* intelligent tutoring systems as AITS, I consider ITSs in general to be capable of adaption and use this term.

An ITS contain four modules: a teaching material module, a learner model module, a tutoring strategy, and an interface (Fig. 1) (Erümit & Çetin, 2020; Nagao, 2019; Nwana, 1990). These modules are described in detail in the 'AI as an intelligent tutor' section.

Most research on ITSs have emphasised its use in traditional educational settings (Erümit & Çetin, 2020; Humble & Mozelius, 2019). I here focus on the use of the principles of an ITS in situations in which human teachers and tutors, and suitable human-produced teaching material, do not exist. This is where AI comes into play, and a brief description of the key technologies involved is now in order.

Key Concepts: AI and Machine Learning

Russell and Norvig (2014) examine a range of definitions of AI ranging from the imitation of human thought processes or action and rational thinking and action. When discussing AI, I refer to computer systems that can perform tasks commonly perceived as requiring intelligence had they been performed by humans (Brundage et al., 2018; Kurzweil, 1990). Since around 1990, the term *artificial intelligence* has declined in

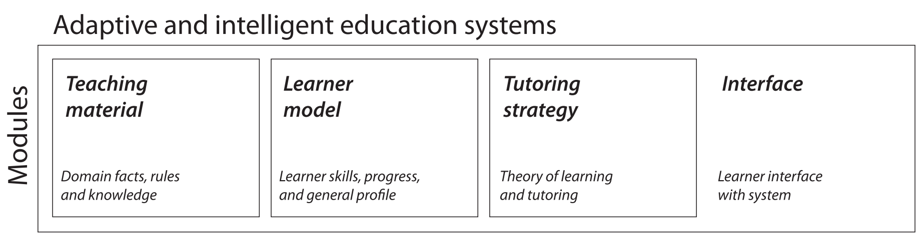

Fig. 1 Generic ITS

popularity, while the terms ‘deep learning’, ‘machine learning’, and ‘neural networks’ are used increasingly often (Kennedy, 2019). These are terms I shortly return to.

I will mainly focus on Google’s DeepMind system, which Google describes as pushing the boundaries of AI for ‘positive impact’ (Google, 2020c). The particular instance of this system that I focus on is the Go and chess playing AlphaZero, which succeeded AlphaGo—the star of a revolution in the world of Go, as we shall shortly see (Google, 2018, 2020a). The DeepMind team has made a number of systems, and they have published a number of articles describing, for example, AlphaGo (Silver et al., 2016), AlphaZero (Silver et al., 2017), the general-purpose game playing MuZero (Schrittwieser et al., 2020), and the more recent AlphaFold, which can predict protein structure (Jumper et al., 2021). DeepMind has also made AlphaStar, which plays the game StarCraft II and managed to achieve the level of grandmaster (Google, 2020b).

I will not go into the technical details of these systems, but will note that AlphaZero is based on what we refer to as machine learning. This involves applying algorithms that ‘learn’ from experience, and the process is closely related to pattern recognition (Bishop, 2006). The kind of machine learning used is called ‘deep reinforcement learning’, which relies on neural networks. The deepness of these networks relates to how many layers they consist of. This allows for unsupervised *reinforcement learning*, where we provide the program with a goal—or target—and constraints and let the system find its own way towards optimal results (Sutton & Barto, 2018).

AlphaGo was originally trained by being fed huge amounts of human games to learn from, but the recent instance of AlphaGo learns by playing against *itself* (Google, 2020a). The system no longer needs humans to demonstrate what is a good move—it can find better moves by trial and error. Machine learning algorithms are not new, but the computing power we now have, combined with ever larger datasets, lets us do new and impressive things with somewhat old tools (Danaher, 2016; Marcus & Davis, 2019).

While older game playing AIs, such as IBM’s Deep Blue, is quite different from the deep learning-based AlphaZero, it is still considered AI. Despite playing ‘brutally and materialistically’, such programs are clearly encompassed in the aforementioned definitions (Strogatz, 2018). Korf (1997) argues that Deep Blue used ‘alpha–beta minimax search with a heuristic static evaluation function’, and that ‘if any technique deserves to be called AI, this one does’. As we have seen, AI encompasses a wide array of phenomena, and the particular technology in use is not of central important for the use of AI for tutoring.

The AI discussed here is not assumed to *understand* anything about the games it plays, its meaning, or anything else (Marcus & Davis, 2019). If we relate the discussion to Bloom’s famous taxonomy, it would not fare well on the cognitive level called understanding, but it could perform exceedingly well at the levels both below and above (Fulbright, 2020). In Fulbright’s (2020) proposed ‘expertise level’, consisting of a range of skills, AI is not assumed to be capable of, for example, understanding or teaching. Furthermore, AI systems are not embodied, and it is consequently worth noting that while, for example, the cognitive paradigm of ecological-enactivism is important for explaining humans’ purposive engagement with their environment (Rolla & Novaes, 2020), it does not necessarily apply to AI systems, which could perhaps be more appropriately described by cognitive theories based on traditional models of information processing, control, storage, and retrieval (Carvalho & Rolla, 2020). However, I will show that the system of *scaffolding functions* based on Vygotsky’s theory might still enable ‘dumb’ AI to function as a form of tutor.

AI as an Intelligent Tutor

AlphaZero is not built as an ITS, but turning it into one might turn out to be one of the most impactful applications of modern-day AI. It can be argued that AlphaZero can create *learning material*, and we will shortly see what this entails. What remains are the modules pertaining to the particular learner, teaching strategies, and an effective learning interface. The interface is not emphasised here, as I focus on how AI can perform functions related to the teaching material module and to a certain degree to the learning model module.

The *teaching material* module is where modern AI plays the key role in my framework. The AI can actually play against the human champions and thus provide learning material and demonstration 'live', and it also produces examples of strategies and tactics at a level no human is capable of. As noted by Nwana (1990), this module can be completely opaque (blackbox), or it can involve more transparent processes and learning material with various levels of accompanying explanations.

A *learner module* contains data on individual learners and is based on learning analytics. What are the strengths and weaknesses of a particular player, and what are their patterns of learning and progress? By analysing the behaviour of a learner, such a module gradually becomes better at tailoring the learning content in the optimal way for each individual (Nagao, 2019). This module could be crucially important for achieving *efficient* learning, and it could also be the crucial aspect of providing tutoring that *motivates.* While some players may be encouraged by strong resistance, others might lose their motivation and give up when faced with the same. One model for analysing each learner could be to map the motivation and learning traits of each learner (Hwang et al., 2004). This will be based on statistics and the analysis of patterns in the learner's behaviour, and no fundamental understanding of humans in general or individuals by these AI systems is assumed.

The role of AI and machine learning in creating detailed personality profiles has received attention far beyond the field of education. Yeung (2017) and Sætra (2019), for example, highlight how Big Data combined with AI and 'nudging' provides potentially very effective means of influencing behaviour. Detailed learner profiles can be combined with the use of AlphaZero's evaluation function to identify strengths and weaknesses in the learner's abilities. In chess and Go, for example, the detailed profiles might relate to activity patterns and data related to motivation. In addition, domain specific profiles, identifying strengths in the various phases of a game (opening, middle game, and endgame in Go, for example) can help determine what form of practice is most effective for each particular learner champion. Similar use of AI is already prevalent in existing adaptive ITS's (Humble & Mozelius, 2019). An ITS will for example begin with easy lessons and gradually map and adapt to the skill, progress, and individual differences between learners, and it can also take account of affective factors such as feelings, emotions, and moods (Alhabbash et al., 2016). It is in the following assumed that an AI-based system can perform important functions related to learning analytics.

The module of *learning strategies* consists of a collection of research-based strategies for tutoring, and when combined with the learner module, it can apply the appropriate strategy for each learner (Nagao, 2019). Hwang et al. (2004) show how learning can be understood as 'learning dynamics, learning energy, learning speed, learning force, and learning acceleration', and that understanding the characteristic of a particular learner, and their learning situation, can enable us to find an optimal learning strategy. Vygotsky's principles of learning are used as the general learning theory in the framework I propose, while domain-specific strategies related to the area of application are developed on the basis of these principles.

Finally, there is the *interface module*. This is the user interface of the AIES, and while it can be a full chess or Go board, it could—and probably should—also involve various other interfaces. One obvious possibility is an interface for solving particular 'problems'. Chess and Go problems are specific isolated parts of the games that allow players to practice various aspects of their games. For example, the 'intelligent consultant system' described by Lazzeri and Heller (1996) provides one such interface for learning proper strategy in the middlegame of chess.

Vygotsky's Theory of Human Learning

A key purpose of this article is to make sense of arguments related to how humans and machines are jointly able to perform better than humans alone (Fulbright, 2020; Heaven, 2019; Kasparov, 2008). Various examples seem to support this idea, but the explanation of how this occurs has received far less attention. This is what I attempt to address by using Vygotsky's theory to explain (a) how humans learn and (b) how AI can perform an important role in human learning processes. This is where modern computer science is connected to the world of developmental psychology and theory of education and human learning—a crucial step for understanding the role of technology in learning (Nkambou et al., 2010; Nwana, 1990).

Vygotsky is a familiar name to those involved in education and developmental psychology. His work is varied, and I here focus specifically on three concepts often related to this psychological theory: the *zone of proximal development (ZPD)*, *scaffolding*, and *more competent others*. These core concepts are considered universal and can be used as the basis for learning in any domain. While this article emphasises chess and Go, the principles discussed are assumed to be relevant to other domains where AI surpasses us as well.

A core component of Vygotsky's theory of development is his emphasis on the affective and contextual aspects of meaningful learning. The basic idea is that people learn in social contexts, and Vygotsky highlights the primacy of the social over the individual level, as all development appear first on the social level and only later on does it appear on the individual level (Vygotsky, 1987). He states that we 'become ourselves through others' and believes it is a general law that development occurs in three stages: development in itself, then for others, and the for oneself (Vygotsky, 1989). The social context, according to this theory, is where we reason and reach new insights in dialogue and interaction with each other and where learning is a joint product several individuals (Ormrod, 2016). When the social context is considered crucial for learning, and when we emphasise the role that society has in individual learning, we often use the term *socio-cultural* theory of learning (Ormrod, 2016). The term is here used to describe the approaches to learning based on Vygotsky's ideas of cognitive development.

According to Vygotsky, people do not simply absorb and store information. We engage with it, and organise it, and we do so in different and at times unique and peculiar ways (Ormrod, 2016). There are important and extensive debates about what the proper interpretation of, and the proper implications for teaching that follow from, Vygotsky's theories (Ormrod, 2016). Some focus on the content of instruction, for example, while others focus on forms of learning and participation (Daniels, 2007). These differences have important implications for how the core concepts often connected to Vygotsky are interpreted

(Daniels, 2007), and this has led to a situation in which many different approaches are said to be based on Vygotsky (del Río & Álvarez, 2007).

Another problem follows from the internal inconsistency in parts of Vygotsky's philosophy, as his ideas were developed over a short period of time and changed during this period (Daniels et al., 2007). For example, throughout his career he shifts from focusing on *external* mediation to being more concerned with *internal* mediation (del Río & Álvarez, 2007). Mediation refers to the idea that we never interact directly with the world, but through signs, which function as the mediational means of accessing an understanding of both physical and social phenomena (Wertsch, 2007). I do not attempt to adjudicate on what is the proper interpretation of Vygotsky and will base the following discussions on mainstream understandings of the core concepts inspired by his psychological theory.

We often hear that it is lonely on top. And more importantly, there is great potential for stagnation up there. When considered in a socio-cultural context, this makes perfect sense because human learning is, to a large degree, based on social contexts and learning from *more competent or knowledgeable others*. Vygotsky (2019) highlights the importance of the presence and participation of a *final form* in learning contexts, such as a developed adult which in a sense demonstrates the final form of language for a developing child. Without the presence and participation of appropriate final forms, development is hindered and potentially arrested (Vygotsky, 2019). This is also highlighted by Cole and Engeström (1993), who argues that the future or mature state is a vital part at all stages of learning. The final form or future state for a human champion has historically been nowhere to be found, but AI potentially constitutes a unique and at least partial form of this kind. Vygotsky (1997) discussed 'The problem of giftedness' in *Educational Psychology*, but he related this more to a general need to provide individual accommodations for all needs and did not discuss the potential of gifted children not being able to find more competent others.

We might also see the danger of stagnation as a natural result of diminishing marginal returns on time spent on a specific subject, but I will focus on the lack of opportunities to learn from others and the absence of final forms and more competent others. In addition, *development* is not synonymous with *learning;* learning is what *creates* the potential for development (del Río & Álvarez, 2007).

The ZPD

However skilled we are, and whatever level we are at in some context, there are certain things that are *just* beyond our current level. Things we do not yet fully understand, but that are so closely connected to what we do know, and close enough in level, that we can *almost* grasp it. This is one way to describe Vygotsky's idea of a ZPD. As a Go player, I have personally reached a certain (low) level of knowledge of the game and its fundamental concepts. At my current level, I have a decent grasp of what goes on in a Go game with an equal, and I can operate on my own. In the ZPD, however, I can only operate successfully under 'guidance or in collaboration with more capable peers' (Vygotsky, 1980, p. 86)—it is the zone of *potential* development or the next zone of development (Kellogg & Veresov, 2019).

In the ZPD, a person is able to perform tasks *with assistance,* and this interpersonal experience can lead to the processing and internalisation of such experience, which is conducive to the construction of intrapersonal experience (Lindblom & Ziemke, 2003). For example, a child can *imitate* what others do in the ZPD, and this can still lead to learning. This can be explained by Kaptelinin's (1996) activity theory and the use of cognitive tools, or *artefacts*, which are devices that ‘maintain, display or operate upon information’ (Norman, 1991, p. 17). AI systems can be said to be such artefacts, and cognitive artefacts are almost identical to the *psychological tools* of activity theory (Engeström, 2015). The learning context or ‘teaching system’ can consequently be understood as ‘stretched across’ both individuals and *things* (Cole & Engeström, 1993), and this is particularly interesting regarding the potential role of AI in learning. Vygotsky also noted that a more competent other did not need to be physically present in the context in which learning was taking place (Daniels, 2007), further opening for non-traditional forms of support in the ZPD.

Kaptelinin (1996) refers to three stages of artefact use that are highly significant for AI’s potential to be our champions’ more competent others. The first stage of tool use is inefficient, as we do not yet master the tool. Performance of the given task will be the same as without the tool, or perhaps even *worse*. In the second phase, however, we have learned to use the tool effectively, and we are thus able to perform the task *better* than we would have done without the tool. The most interesting phase is the third. Here, Kaptelinin (1996) states, we can perform the task more effectively than at the outset, even without using the tool.

Stagnation, then, is a natural consequence of learning activities that are not performed in the ZPD; when learning activities are performed at a level we already master, learning is inefficient (del Río & Álvarez, 2007). While various tutors and tools based on existing knowledge may help champions improve certain aspects of their capabilities, it is exceedingly difficult to find teachers or tutors that can effectively operate in the ZPD of champions. They are, after all, the most competent of all humans in their domains. It is also important to note that the *other* person involved in learning need not be an instructor—they can also be a peer. In chess, good players often cooperate in teams, and peers are often also hired as tutors or teachers. Garry Kasparov, for example, has been involved in the teaching of Magnus Carlsen, and Peter Heine Nielsen, Carlsen’s current trainer, is also a Grandmaster. I will soon return to the idea of the *more competent other*, and the problems of finding these for champions, but first some basics about the process of *scaffolding*.

Scaffolding

In the ZPD, we can perform successfully with *support*. This support is often referred to as *scaffolding*. The term is now used in a wide variety of ways, and Daniels (2007) warns against the danger that it will be appropriated by proponents of theories far removed from Vygotsky’s original theory if we use it for just about any practice of assistance in learning. I use the term to refer to an approach to learning in which the role of the learner, and not what is being learnt, is simplified, and in which ‘the overall emphasis is on the creation of a pedagogic context in which combined teacher and learner effort results in a successful outcome’ (Daniels, 2007). It might be noted that learning may also be *self-managed*, but even such self-managed learning can be scaffolded (Lesgold, 2019). This could be the case of a chess player using an expert chess application for inspiration, but having a more capable peer helping them interpret the actions of the programme and relating the techniques and new ideas to the learner’s current level of understanding. This also highlights an asymmetry between AI systems and humans, as we have already seen how systems such as AlphaZero achieve superhuman levels of skill by simply playing against itself. Humans arguably need others for development and would not benefit greatly from simply playing chess against themselves without any support or guidance. This is particularly relevant in instances involving higher functions and not the perfection of, for example, motor skills such as hitting a particular tennis shot.

There are certain characteristics that are required for effective scaffolding, and Wood et al. (1976) highlight six key ‘scaffolding functions’: *Recruitment, direction maintenance, frustration control, simplification, marking critical features,* and *demonstration*. It is notable that this typology of tutoring functions frees us from considering *understanding* and *teaching* as skills required for tutors. What Fulbright (2020) refer to as basic skills are here broken down into components that show that AI could be able to perform teaching functions and foster understanding, without itself understanding anything.

In order to determine the potential of AI tutoring, we must examine to what extent a machine can perform each of these scaffolding functions. I divide these functions into *motivational* functions (the first three) and *instructional* functions (the latter three). Haake et al. (2015) use the terms *cognitive* scaffolding and *motivational* scaffolding, but relate these more to the context of a situation than to the functions I described above, which are attributable to a more competent other.

Motivational Functions *Recruitment* involves getting the learner interested in the task at hand (Lesgold, 2019). This is very important for tutoring children, for example, but we should not underestimate the role of motivation and ensuring the learner’s ‘interest in and adherence to the requirements of the task’ even for adult champions (Wood et al., 1976, p. 98). Key factors in learning is *interest* and *attitude*, and fostering interest and an attitude conducive to learning are important aspects of recruitment (Hwang et al., 2004).

Next, we have *direction maintenance*, and this involves making sure the learner stays focused on the specific task (Lesgold, 2019). This is particularly important when involved in challenging tasks that the learner does not yet master, as this can lead to a tendency to *regress* and turn to tasks and behaviours that the learner has already mastered. Direction maintenance involves providing encouragement in terms of positive feedback, particularly in order to make sure that the risk-taking involved in attempting new tasks is experienced as worthwhile (Wood et al., 1976).

The last of the motivational functions is *frustration control*, and this ties into the previous function. When new tasks are attempted, there is bound to be some degree of failure in the initial phases, and the tutor’s role is to both dampen the feeling of defeat and make sure to abort or change specific activities before frustration becomes debilitating (Wood et al., 1976).

Instructional Functions The first of the instructional functions is *reduction of degrees of freedom*. Simplifying the task at hand, so to speak (Lesgold, 2019). When we face a task we do not master, we initially experience a frustratingly large amount of seemingly plausible ways to solve the task. A more competent other who understands the task better than us can perform parts of the task, while we are left to perform a subset of the task. In the frame of affordances, a more competent other might help a learner identify the relevant affordances from the richer landscape of affordances, as proposed in the skilled intentionality framework of Rietveld et al. (2018), building on the ecological-enactive paradigm of cognition.

An important part of scaffolding, then, is breaking up complex tasks into sets of simpler operations (Mubin et al., 2013). In Go, for example, the game could be considered in term of whole board opening play (*fuseki*), local opening plays (*joseki*), midgame (involving both global and local strategies and tactics), and endgame (*yose*).

Next, there is the function of *marking critical features*. As with the previous function, the complexity of something we do not master can be overwhelming. The more competent other can help us immensely by simply pointing out what factors are central to the task. For example, if the learner attempts to imitate the actions of a master at a task, the tutor can point out the central missing aspects in the learner's work (Wood et al., 1976). Key moves in a game can also be identified by current AI by using their evaluation functions. AI plays by consistently evaluating the status of the game and the relative merit of various positions and plays, and can identify and mark moves that significantly worsen the champions position, for example. This is a crucial part of what enables AI to learn much more effectively from playing itself than humans do.

Finally, the tutor can perform *demonstration*. More than simply performing the task the way they would normally do it on their own, this involves giving an 'idealised' version of the performance they want the learner to learn (Wood et al., 1976). According to Lesgold (2019), this involves having the learner '*compare* their task performance to an *expert version*'.

AI-Based Expert Scaffolding System

The system I propose is an AI-based expert scaffolding system. The technical details of the system are not developed here, and the emphasis is on how the various functions of an ITS, combined with modern AI and Vygotsky's theory of learning, can explain how AI can lead human champions towards new levels of knowledge.

The system in Fig. 2 shows how a specialised AI system, such as AlphaZero, provides the material for the teaching material module, while being involved in the analysis performed in the learning module, while also drawing on these results in order to provide appropriate learning materials for each student. The tutoring strategy in this system is based on Vygotsky's ideas about the ZPD and scaffolding. This allows tutoring to occur through the provision of demonstration and instruction materials of a superhuman level, supported by the provision of scaffolding. A key challenge in this system is that current AI does not have the abilities of understanding or teaching, which invites the possibility that human experts should also be involved in the process. These possibilities are developed in more detail in "Discussion: AI as the More Competent Other", after an examination of how AI operates in the ZPD of human champions in the games of chess and Go.

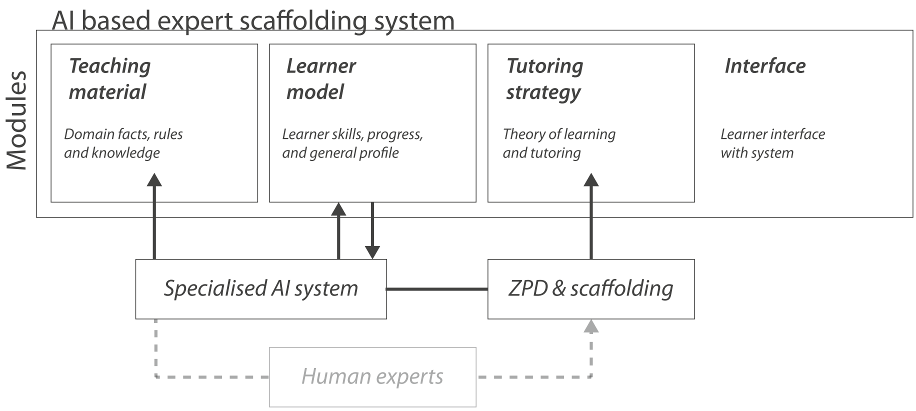

Fig. 2 AI-based expert scaffolding system

Dethroned Humans and New Competent Others

The main cases I examine are the games of chess and Go. These two games are domains in which AI has demonstrated superhuman abilities, and in which there is much debate about the consequences of such AI supremacy. Chess and Go are examples used to demonstrate the feasibility of using AI as the more competent other and to examine whether or not human champions actually perceive AI in ways compatible with what is required for the proposed framework to have any validity.

While I focus on these examples, the mechanisms involved in scaffolding and the ZPD are assumed to be universally applicable. Energy use, science, and medicine are examples of new domains in which the DeepMind team explores the application of its AI (Heaven, 2019; Strogatz, 2018). If similar progress is made in these domains, the lessons from chess and Go here described will in principle be applicable also there. It is also worth noting that DeepMind's latest conquest in the gaming world, *StarCraft II* provides an example very similar to the ones discussed in relation to chess and Go below (Heaven, 2019).

Chess and Go

Games, you may say, are frivolous cases for highlighting the potential for human intellectual progress. If you do, I would argue that games are serious business. Particularly when it comes to the games of *Go* and *chess*.

Go is the Japanese name for an ancient game of strategy that is hugely popular, particularly in Japan, China, and Korea. *Wei'chi* is the Chinese name, and throughout Chinese history, the game has had a big impact. A past-time for some, but much more for others, as it 'has been a favourite game of strategy of Chinese generals, statesmen, and literati from the former Han dynasty (206 B.C.-8 A.D.) to the times of Mao Tse-tung' (Boorman, 1969).

The games of Go and chess are even used to symbolise the difference between the strategic approaches of the west and the east. Kissinger (2011), in his book *On China*, states that chess symbolises the west's desire for the 'decisive clash of forces', while the Chinese game of wei'chi symbolises the 'subtlety, indirection, and the patient accumulation of relative advantage' more prevalent in eastern strategic thinking. Chess, he states 'is about the decisive battle', while Go 'is about the protracted campaign' (Kissinger, 2011). Finally, Kissinger (2011) thoroughly reveals his favourite as he states that 'chess produces single-mindedness; [Go] generates strategic flexibility'. Regardless of how we see the two games, they are considered to be of great importance. Seeing how the US and the USSR battled for supremacy on the chess board only serves to further strengthen this view (Bernstein, 2012).

Computer chess was one of the original problems one sought to solve with the development in AI (Korf, 1997; Shannon, 1950). In 1997, AI achieved a huge success when it beat the champion Garry Kasparov (Campbell et al., 2002). IBM was the company behind this success, and the application was named Deep Blue. It is also interesting to note that Kasparov had only five years earlier 'scorned the pathetic state of computer chess' (Kurzweil, 2015, p. 148). It was not pathetic in 1997, and it has become increasing less so ever since. While DeepMind was the first and most famous chess-playing system, Stockfish has gotten most press in recent years, while AlphaZero today is regarded as the most capable system (van den Herik, 2018).

With chess being conquered by AI, a lot of people placed their hope in the game of Go being able to withstand AI's assault. When Deep Blue beat Kasparov in 1997, I, as a mediocre Go player, was still able to beat the best Go programs. The brute force approach of DeepMind did not accomplish much in this game with far more variations, and it was not until 2016, nearly *20* years later, that AlphaGo managed to beat a human champion and become recognised as superhuman in skill. Lee Sedol, a Korean top professional, lost against Google's new creation (Chouard, 2016).

Chess and Go are games of huge impact, and much human pride has been built, and destroyed, over chess and Go boards. No wonder, then, that we cared so deeply when machines surpassed us and became the new masters of what was once the pinnacle of the human intellect. But how do the players relate to their new masters—with indifference, contempt, or admiration?

Before moving on to how AI can help our champions, I will briefly show that our professionals and champions of these games *do* care and that AI *does* have the potential to teach our champions new things.

AlphaGo and Lee Sedol

> *'It's not a human move. I've never seen a human play this move,' he [Fan Hui] says. 'So beautiful.' It's a word he keeps repeating. Beautiful. Beautiful. Beautiful.* (Metz, 2016)

The best of five games match between Lee Sedol and AlphaGo in 2016 was a big event. Later on, in the chess world, AlphaZero burst onto the scene by beating another computer (Stockfish) already considered the superior of human champions. In the world of Go, however, AlphaGo faced the human champion, who was in fact considered to be better than any computer at the time. When AlphaGo managed to win, *convincingly*, it was a dramatic event, both in the game of Go, AI, and human intellectual history. In terms of the demonstration function of scaffolding, AI definitely proved itself capable.

In order to see how human professionals perceived this new phenomenon—AlphaGo—let us see how Lee Sedol himself reacted to this game. First of all, he stated that it made him question the traditional approaches to the game itself (Economist, 2016). This is, as we shall see, similar to what occurred when AlphaZero later conquered chess anew.

Traditions and classical approaches to games such as chess and Go are very important, and this is natural, given that no human being is able to fully comprehend the games. Thus, generally accepted 'best practice' ways to play become the guideline. It is interesting that AlphaGo at times supports such traditions, as they are often the best ways to play, but that it happily breaks from them whenever it has found better plays. While it is arguably absurd to attribute intentions and a will to AI, we see that quite a few *do* when they encounter AlphaZero. Strogatz (2018) compares AlphaZero to 'brutes' like Deep Blue and states that AlphaZero 'seemed to express insight', that it played 'intuitively', with a 'romantic attacking style'. He goes on to explain how AlphaZero was *toying* with Stockfish, like a matador. Heaven (2019) similarly notes that DeepMind's software displays 'what looks very much like creativity and – whisper it – intuition'. While I will not discuss machine agency and anthropomorphism here, it is interesting to note how modern AI inspires such perceptions and descriptions, as it relates to its motivational scaffolding functions. This is also important for understanding how AI can be construed as *other* even by those, like me, who believe that AI systems are severely limited when it comes to abilities of being meaningful social partners capable or reciprocal relationships (Sætra, 2020, 2021). For the sake of the current discussion, it might be sufficient that learners experience these systems are meaningful others, and this finds support in the relational turn in the field of robot ethics (Coeckelbergh, 2010; Gerdes, 2016; Gunkel, 2018).

Fan Hui, a professional Go player, when commenting on another AlphaGo game against the top professional Ke Jie in 2017, refers to AlphaGo's 'own unique philosophy':

> *AlphaGo's way is not to make territory here or there, but to place every stone in a position where it will be most useful. This is the true theory of Go: not 'what do I want to build?', but rather 'how can I use every stone to its full potential?'* (DeepMind, 2020)

It is these moments of breaking with tradition that baffle the professionals. At the beginning of the game against Lee Sedol, many top professionals initially perceived some of AlphaGo's moves as mistakes, only to be taken aback when they realised that the AI had simply understood far more than them, and the seeming mistakes were in fact brilliant moves.

The word *beautiful* is often used to describe AlphaGo's play, and the professional Lian Xiao states that AlphaGo 'could actually broaden the horizon of Go playing' and 'bring more imagination into Go' (DeepMind, 2020). Another top professional, Gu Li, agrees and says that AlphaGo's games 'will give us new ideas about how to play. If there's an opportunity, we should do more of these' (DeepMind, 2020). AlphaGo here seems to both inspire professionals to learn and to demonstrate new and novel ways of playing the game effectively.

Ke Jie himself, after losing three games to AlphaGo, states that it is 'perfect, it's just flawless, merciless. … I don't think I could catch up with it in my lifetime' (Yun, 2017). And this, unfortunately, takes us to one major challenge regarding the rise of the superhuman machines. I mentioned the potential consequences of superhuman AI in the introduction, and in the world of Go, we have seen one unfortunate consequence directly related to the recruitment function of scaffolding. Lee Sedol, the face of the human defeat to AI in Go, recently decided to resign from professional play. Why? Because he considers AI to be unbeatable, and instead of encouraging a chase, this seems to have drained his motivation:

> *'With the debut of AI in Go games, I've realized that I'm not at the top even if I become the No. 1 through frantic efforts,' Lee said. 'Even if I become the No. 1, there is an entity that cannot be defeated.'* (Webb, 2019)

Beautiful, innovative, and inspiring. However, it is also *discouraging* to some, and this matters for the potential of an AI such as AlphaGo to be a tutor for champions.

Champions of Chess and AlphaZero

> *Through learning about AlphaZero we can harness the new insights that AI has uncovered in our wonderful game of chess and use them to build on and enhance our human knowledge and skills. (*Sadler & Regan, 2019*)*

What, then, has happened in the world of chess? The current undisputed champion of chess is Magnus Carlsen, so who better to turn to than his trainer Peter Heine Nielsen, who wrote an article about the meeting between AlphaZero and Carlsen. Nielsen (2019) compares the release of AlphaZero in 2017, and how it beats Stockfish, to the arrival of a spaceship landing in the middle of London. The following shock was not only inspired by awe, but also by the fear that opponents would be the first to take advantage of this technological breakthrough. The games were 'beautiful', and there were 'interesting novelties' in them (Nielsen, 2019). AlphaZero not only beats Stockfish, but it did so through brand new ways of playing the game, including sacrifices previously thought to be madness against an opponent such as Stockfish (Nielsen, 2019) Following this breakthrough, DeepMind members also released an article on the algorithm behind it and how AlphaZero is one step away from becoming a *general* game-playing AI (Silver et al., 2018).

Nielsen (2019) even states that AlphaZero had *renewed* chess, but that, for quite some time, players were left to do their own interpretations of just how it had achieved this. A new champion had arrived, but it was a champion that did not teach, or explain. Nielsen states that the situation improved with the arrival of the book *Game Changer* in 2019 (Sadler & Regan, 2019). The authors had access to a lot more games than had been available previously and were able to get experiments run in order to develop a more thorough understanding of how AlphaZero 'thinks'.

Of particular interest is the notion that AlphaZero has somehow *changed* Magnus Carlsen. Nielsen (2019) writes of how Carlsen early in 2019 stated that he had changed his playing style, 'acknowledging inspiration from AlphaZero, and Daniil Dubov'. Magnus Carlsen's father, Henrik Carlsen, agrees that Magnus's style has changed and that he has become more aggressive (Strøm, 2019). Henrik Carlsen also states that AlphaGo 'has been an important inspiration. It is something brand new …. Not just a computer program which concludes, but more like a fantastic human being. A mix of human and machine' (Strøm, 2019). Leontxo Garcia is also quoted as saying that he believes AlphaZero has had a 'very positive influence' on Carlsen, seemingly clearly performing motivational functions. Magnus himself states:

> *AlphaZero has been important for a lot of people, me included. It has been interesting to see those games. I do not try to play like it – I largely play as I have done before – but there may be some new opening ideas.* (Strøm, 2019)

Nielsen (2019) goes on to venture a comparison of the *mindset* of Carlsen and AlphaZero:

> *No matter whether efficiency means sacrificing material or being incredibly solid, they will do what is required! Regardless of whether they are the pinnacle of AI or the Chess World Champion.*

Sadler and Regan (2019) suggest that we can use AlphaZero to 'enhance our human knowledge and skills'. Garry Kasparov (2019), a man with experience both of being dethroned by a computer and training Magnus Carlsen, writes about *AlphaZero and the Knowledge Revolution.* He also speaks of working on the 'potential of human plus machine' (Kasparov, 2019). He notes that the approach of AlphaZero is very different from the brute force approach of Deep Blue that defeated him long ago:

> *AlphaZero isn't just applying human knowledge and plowing through billions of positions to generate moves – it's creating its own knowledge first. And, based on its results and my observations, the knowledge it generates for itself is unique and superior. We aren't just getting faster results the way we do from a calculator. Instead of a postcard from a far-off land, it's a telescope that has the potential to let us see for ourselves.* (Kasparov, 2019)

If AI can indeed make us 'see for ourselves' instead of merely providing us with unfathomable results, a case can certainly be made that AI is now performing a number of scaffolding functions.

Paranoia and prejudice are natural reactions to machines that surpass us, Kasparov (2019) states, but he also says that we must overcome these reactions. This is the 'next phase of human–machine cooperation, to accept that machine knowledge and judgment can be superior to our own. Instead of just using machines as tools—the "centaur" model—the machines become the experts and humans will oversee them, I call it a "shepherd" model' (Kasparov, 2019). In the introduction to *Game Changer,* Demis Hassabis, founder and CEO of DeepMind, states that his hope 'is that the games and analysis in this book will help to spark a new era of creativity in chess, and that players will not only incorporate some of these ideas into their own games, but also be inspired to find new styles of their own' (Sadler & Regan, 2019).

If humans are the shepherds of artificial intelligence, this implies that humans are really in control after all. I propose a different perspective on this new relationship, and that is one where the *AI* is in fact the shepherd, so to speak. It is AI that is in lead—the more competent other of sorts—and it can drag our champions along with it into brand new ZPDs.

Discussion: AI as the More Competent Other

It is now time to examine the validity of arguing that AIs such as AlphaZero can function as the more competent other of human champions and perform some—or all—of the scaffolding functions that are central to Vygotsky's theory.

With these functions in mind, what role might AI have in performing the role of the more competent other—or the *tutor?* I will here consider what roles existing programmes like AlphaZero can perform and will then comment on the possibilities of building more advanced tutors from this.

The role of communication in scaffolding is of great importance, but it is one that is not as relevant to the current state of AlphaZero. It is, however, possible to foresee a future where the machines attempt to explain and communicate their processes for arriving at good decisions. A lot of effort is being expended on the quest for AI that can *explain* its decisions, and pressure is mounting to achieve this (Wachter et al., 2017). If we could understand the decisions of AI systems, this would allow us to both trust and manage them in a way that is different from the current situation (Gunning, 2017). Human–machine communication is a distinct field of study (Guzman, 2018), and recent debates about current large language models such as GPT-3 (Floridi & Chiriatti, 2020) and LaMDA (Google, 2021) serve to show how machines are now communicating at a level that challenges many people's notions of what things such as communication, life, and consciousness are. The latter was made abundantly clear as Google engineer Blake Lemoine stated that he believed their LaMDA system is conscious and a *person* that can or should not be Google's property (Levy, 2022).

While AlphaZero *seems* to have grasped some key insights of chess and Go, these insights are largely unavailable to us without a process such as the one proposed in this article, and it is not immediately obvious that a large language model approach to communication can help close this gap. Human beings want more than *answers*—we want reasons—and Strogatz (2018) discusses a hypothetical 'AlphaInfinity' that would able to directly explain and help us learn from it. Explainable AI would be far better equipped to be a useful more competent other in a scaffolding situation, but in lieu of such developments, the ITS approach here proposed still provides a way to learn from AlphaZero.

AlphaZero playing chess and Go is not explainable AI, but as Robbins (2019) notes, the effort spent on achieving this might very well be misdirected because 'a principle of explicability for AI makes the use of AI redundant'. The reasoning behind such a stance might be that if we could be made to understand how AI reaches its output, we would also be able to do it ourselves. I partially disagree with such a stance, and as shown in "Dethroned Humans and New Competent Others" above, top players *do* benefit from insight into how the programs 'think', even if the thought process is not fully fathomable. While such a situation might not represent the ideal according to Vygotsky's theory, it adheres to some key principles related to the ZPD and scaffolding and arguably provides opportunities for learning that would not otherwise exist.

By considering the functions involved in scaffolding of champions in the games of Go and chess, I show in the following table what functions AlphaZero can perform, what AlphaZero combined with a hypothetical explainable element (X-AlphaZero) might be able to perform, how X-AlphaZero developed with the four modules of an intelligent tutoring system (AlphaZero ITS) might perform, and, lastly, what AI in combination with human beings might perform.

In Table 1 the lower-case *v* indicates that the function can be performed to a certain degree, while a capital *V* indicates that the task can be performed more or less fully.

AlphaZero as it currently exists can partly perform the functions of demonstration, reduction of complexity, and marking critical features. It does not do so fully, and achieving these effects requires much work on the part of the learner. However, AlphaZero and similar applications are continuously monitoring the situation on the game board, and the relative merit of available moves. It can, for example, easily identify which moves are good or bad and which moves it would itself have played. While it cannot really explain why, this goes a long way towards marking critical features.

Also, AlphaZero is not actively performing any motivational functions of scaffolding. However, the mere existence of an entity that surpasses the human champion *can* be motivational, and one interesting aspect of how AlphaZero might today perform tutoring functions involves a situation in which a champion mimics the play of AlphaZero and in turn develops into a stronger player. According to Kaptelinin (1996), this occurs through a process of internalisation, by performing certain acts without fully understanding them and only gradually mastering and grasping the reasons behind them.

If we had explainable AI, this would improve the instructional scaffolding functions performed by AlphaGo by providing easier access to the processes and mechanisms behind its play. One example, already partly available, is the aforementioned insight into the continuous evaluation of all available moves. AlphaZero could, in addition to marking critical game changing moves, for example, provide a map of the board with various colour coded maps to show the areas of most interest, those that are most pressing, and even what sort of end states it sees as likely when considering various moves. This is not fully explainable AI, but it does provide non-trivial assistance in learning from its demonstration.

However, such explainability would perhaps do little for the motivational factors. This is where AlphaZero combined with the three missing modules of a proper ITS

Table 1 Various forms of tutors and the functions of scaffolding

Function	AlphaZero	X-AlphaZero	AlphaZero ITS	AlphaZero + human
Recruitment			v	V
Direction maintenance			v	V
Frustration control			v	V
Simplification	v	V	V	V
Marking critical features	v	V	V	V
Demonstration	v	V	V	V

could break new ground (the learning material would be produced by the AlphaZero AI). By having a learner module, learning strategies, and an effective interface module, AlphaZero ITS could perform *all* of the function in scaffolding, even if there might be certain limitations in how successfully it performs the motivational functions. However, *affective* AI and social AI of various sorts might here be used in order to further improve AIs' scaffolding abilities (Sætra, 2020). Affective AI here simply refers to AI that is used to detect and potentially respond to human emotions and not some deeper affective states in the AI systems themselves.

It is worth noting that I have added one column full of capital *V*s, and this is the column where one or more human beings are partnered up with AlphaZero in order to perform the scaffolding *together*. This is what is being done by Magnus Carlsen today, and it seems to be working. A team of experts interpret and expound on the instructional data provided by AlphaZero, and, doing what humans do best, they perform the motivational functions better than any current AI could do. Until AlphaZero and similar programmes are built as complete tutoring system, this might be the ideal model for champion tutoring. While AlphaZero plays its best games without human assistance, when viewed from the *human* perspective, Kasparov's (2008) statement that 'man plus machine is stronger than either' seems correct. By using such a team-based approach to learning from AlphaZero, there is an even greater chance of reaping the benefits involved in the social learning described by Vygotsky.

This connects to the previously mentioned concept of distributed cognition. The AI powered education system here proposed can also, partly by accident, lead to the realisation of some of the benefits of what Chou et al. (2003) call *learning companion systems*. Such systems allow learners to learn from *teaching* in addition to learning, and they are designed to motivate the learner to process information in order to convey it to another—thus engaging in 'learning by teaching' (Chou et al., 2003). This is also a key aspect of Vygotsky's theory of learning, and one key benefit in the framework I propose is that human champions might learn more than they could learn from the AI alone if they are also involved in explaining, for example, why a particular move made by the AI is good to their team of teachers.

However, even if groups are formed, there is no guarantee that working in a group fosters truly collaborative work and scaffolding (Nussbaum et al., 2009). This is also a skill that must be learned, and actively nurtured, which then becomes a central task for the leaders of such teams. That scaffolding is effective in small groups is demonstrated by Molenaar et al. (2010), and they also show that the type of scaffold that is performed matters.

The interaction between humans and computers in such constellations is an important field of study of HCI, and activity theory focuses on the *integration of computer tools into the structure of human activity* (Kaptelinin, 1992). Kaptelinin (1996) himself notes, however, that the use of computers by *groups* is one of the limitations of activity theory.

Areas for Further Research

Before closing, there are certain additional factors that are interesting in the extension of the arguments I make regarding the potential for AI to be the more competent other. These include the potential of using more advanced interfaces and *embodied* and *social* AI in order to further improve the scaffolding potential of AI.

For example, Wertsch (2002) suggests that asynchronous communication systems introduce changes that have consequences for 'cognition, identity, and collaboration'. This implies that any AI that can, for example, *speak* and communicate more naturally might have an advantage as a tutor. At the very least, it would better resemble the traditional human–human relationship. The importance of dialogue in learning is analysed by Draper and Anderson (1991), both from a Vygotskian cultural psychological perspective and from a Wittgensteinian perspective.

The potential for *robots*—embodied AI—to create more traditional social contexts that resemble human–human tutoring relationships is particularly interesting. While phenomena such as *joint attention* can be argued to be possible with traditional computer-based AI, humans respond to social cues, and mechanisms such as gaze following are central to achieving joint attention (Okumura et al., 2013; Reid & Striano, 2005). Such mechanisms are much more easily triggered by robots than by screen-based AI. Joint attention is an important aspect of the referential triangle described by Tomasello (2009), which describes how a child, adult, and some object form a triangle where the child learns to understand the function and purpose of the object through the interaction with the adult.

In such situations, we could achieve what Charisi et al. (2015) call *symbiotic interaction*. This involves a 'dynamic process of working towards a common goal by responding and adapting to a partner's actions, while affording your partner to do the same' (Charisi et al., 2015). Signalling intent is central for achieving such interaction, and common ground and mutual trust can be created if such *interpredictability*—mutual signalling and *effective interpretation of signals*—is present (Charisi et al., 2015). Such considerations are of interest because a *social robot* tutor might be even more effective at performing the motivational tasks involved in tutoring. Mubin et al. (2013) discuss the use of robots in education in more detail.

These aspects of AI tutoring relate to the *interface module* of an adaptive and intelligent education system. An advanced social robot would have great potential for employing most of the tutoring functions by not being reliant on a computer and by being able to elicit human social responses. These are both important for making the learning itself more effective through, for example, joint attention, but also through the motivational aspects of perceived social presence.

Another consideration relates to how Vygotsky emphasises the role of the ZPD for being able to *learn* the way humans do and that only humans have the ZPD (Lindblom & Ziemke, 2003). This emphasis on the social nature of learning is of great relevance for the quest for *general artificial intelligence* (Lindblom & Ziemke, 2003). While machines have come far, they are not yet close to the human process of development, and Vygotsky might argue that this is because 'human learning presupposes a specific social nature and a process by which children grow into the intellectual life of those around them' (Vygotsky, 1980, p. 88).

Conclusion

When machines surpass human beings, they evoke a certain degree of fear, resentment, and uncertainty. However, they also free us from existing restraints and create new possibilities. One such possibility is the notion that when AI surpasses us intellectually, it can drag us along with it.

The mechanism I have described involves drawing on Vygotsky's theory to develop an AI-powered expert scaffolding system for human champions. The ZPD is the zone where we all learn new things every day by interacting with and being aided by more competent others. The scaffolding we receive there might be institutionalised and intentional, but many of the aspects of tutoring can also be performed by more competent others without our learning being their goal. By demonstrating tasks and skills we do not master, they aid us in our own quest for mastery. If they effectively perform all six functions of scaffolding, they can tutor the most effectively, but partial scaffolding is also possible.

I have argued that even simple expert AI, such as AlphaZero, can perform certain scaffolding functions for our champions. These are related to demonstration, reduction of complexity, and marking critical features. AlphaZero is not, as of now, performing the motivational functions of scaffolding. But could AlphaZero do more and perform better at what it already does?.

Yes. *Explainable* AI would help improve the instructional functions of scaffolding by providing easier access to the processes and mechanisms behind the level of mastery AlphaZero shows. However, it would do little for the motivational factors. This, I argue, is where AlphaZero combined with the three missing modules of a proper ITS could break into new areas. By having a learning module, learning strategies, and an effective interface module, AlphaZero ITS could perform *all* function in scaffolding, even if there might be certain limitations in how successfully it performs the motivational functions.

While we lack a fully developed educational system for AlphaZero, we do have something that could arguably be even better. AlphaZero combined with a team of human tutors. When a group of human experts together interpret and extend the instructional functions of AlphaZero, and human tutors perform the motivational functions in effective scaffolding, we *do* have a model for effective tutoring of our champions. This enables human champions to experience the full width of social learning envisioned by Vygotsky, where both distributed cognition and the ability to learn from teaching become integral features of the system. This is the model being built by Magnus Carlsen's team, and it has led to a change in the way the champion plays, and those close to him believe that it has contributed to him playing the best chess he has ever played.

These considerations of how AI might perform the function of the most competent other have two immediate implications. Firstly, it provides a model for developing better educational AI. Secondly, it provides a way to understand the hybrid human/machine model of tutoring that is partly being employed already.

AI surpasses us, and many speak of how this will lead to human improvement. In this article, I show how we might understand the processes involved, and that we should develop AI in ways that enables it to drag our champions along with it. In closing, I will add that the implications of this are that we *all* learn, as mastery in certain intellectual pursuits is partly a trickle-down affair, as the improvement of human champions means that there will exist new ZPDs for the second-best human players, and so on. When our champions break new ground, those close to them learn and improve, and in a cascade of learning, so do most of those who are interested.

Funding Open Access funding provided by Ostfold University College.

Declarations

Conflict of Interest The author declares no competing interests.

References

Alhabbash, M. I., Mahdi, A. O., & Naser, S. S. A. (2016). An intelligent tutoring system for teaching grammar English tenses. *European Academic Research, IV*(9).

Bernstein, S. (2012). Valedictorians of the Soviet school: Professionalization and the impact of war in Soviet chess. *Kritika: Explorations in Russian and Eurasian History, 13*(2), 395–418.

Bishop, C. M. (2006). *Pattern recognition and machine learning*: Springer.

Boorman, S. A. (1969). *The protracted game: A wei-ch'i interpretation of Maoist revolutionary strategy*: Oxford University Press, USA.

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... Filar, B. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. arXiv preprint arXiv:1802.07228

Brusilovsky, P. (1999). Adaptive and intelligent technologies for web-based eduction. *Ki, 13*(4), 19–25.

Campbell, M., Hoane, A. J., Jr., & Hsu, F.-H. (2002). Deep Blue. *Artificial Intelligence, 134*(1–2), 57–83.

Carr, N. (2020). *The shallows: What the Internet is doing to our brains*: WW Norton & Company.

Carvalho, E. M. D., & Rolla, G. (2020). An enactive-ecological approach to information and uncertainty. *Frontiers in Psychology, 11*, 588. https://doi.org/10.3389/fpsyg.2020.00588

Charisi, V., Davison, D., Wijnen, F., Van Der Meij, J., Reidsma, D., Prescott, T., ... Evers, V. (2015). *Towards a child-robot symbiotic co-development: A theoretical approach.* Paper presented at the AISB Convention 2015.

Chou, C.-Y., Chan, T.-W., & Lin, C.-J. (2003). Redefining the learning companion: The past, present, and future of educational agents. *Computers & Education, 40*(3), 255–269.

Chouard, T. (2016). The Go files: AI computer wraps up 4–1 victory against human champion. *Nature News*.

Coeckelbergh, M. (2010). Robot rights? Towards a social-relational justification of moral consideration. *Ethics and Information Technology, 12*(3), 209–221.

Cole, M., & Engeström, Y. (1993). A cultural-historical approach to distributed cognition. *Distributed cognitions: Psychological and educational considerations*, 1–46.

Danaher, J. (2016). The threat of algocracy: Reality, resistance and accommodation. *Philosophy & Technology, 29*(3), 245–268.

Danaher, J. (2019). *Automation and Utopia: Human flourishing in a world without work*: Harvard University Press.

Daniels, H. (2007). Pedagogy. In H. Daniels, M. Cole, & J. V. Wertsch (Eds.), *The Cambridge companion to Vygotsky*. Cambridge: Cambridge University Press.

Daniels, H., Cole, M., & Wertsch, J. V. (2007). Editors' introduction. In H. Daniels, M. Cole, & J. V. Wertsch (Eds.), *The Cambridge companion to Vygotsky*. Cambridge: Cambridge University Press.

DeepMind. (2020). AlphaGo China. Retrieved from https://deepmind.com/alphago-china. Accessed 11 October 2021.

del Río, P., & Álvarez, A. (2007). Inside and outside the zone of proximal development. An ecofunctional Reading of Vygotsy. In H. Daniels, M. Cole, & J. V. Wertsch (Eds.), *The Cambridge Companion to Vygotsky*. Cambridge: Cambridge University Press.

Draper, S. W., & Anderson, A. (1991). The significance of dialogue in learning and observing learning. *Computers & Education, 17*(1), 93–107.

Economist, T. (2016, March 19). Artificial intelligence and Go: A game-changing result. *The Economist*. Retrieved from https://www.economist.com/science-and-technology/2016/03/19/a-game-changing-result. Accessed 8 October 2020.

Engeström, Y. (2015). *Learning by expanding*: Cambridge University Press.

Erümit, A. K., & Çetin, İ. (2020). Design framework of adaptive intelligent tutoring systems. *Education and Information Technologies*, 1–24.

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*. https://doi.org/10.1007/s11023-020-09548-1
Fulbright, R. (2020). *The expertise level.* Paper presented at the International Conference on Human-Computer Interaction.
Gerdes, A. (2016). The issue of moral consideration in robot ethics. *ACM Sigcas Computers and Society, 45*(3), 274–279.
Google. (2018). AlphaZero: Shedding new light on the grand games of chess, shogi and Go. Retrieved from https://deepmind.com/blog/article/alphazero-shedding-new-light-grand-games-chess-shogi-and-go. Accessed 8 October 2020.
Google. (2020a). AlphaGo. Retrieved from https://deepmind.com/research/case-studies/alphago-the-story-so-far. Accessed 8 October 2020.
Google. (2020b). AlphaStar: Mastering the real-time strategy game StarCraft II. Retrieved from https://deepmind.com/blog/article/alphastar-mastering-real-time-strategy-game-starcraft-ii. Accessed 8 October 2020.
Google. (2020c). Solve intelligence. Use it to make the world a better place. Retrieved from https://deepmind.com/about/. Accessed 8 October 2020.
Google. (2021). LaMDA: Our breakthrough conversation technology. Retrieved from https://blog.google/technology/ai/lamda/. Accessed 9 July 2020.
Gunkel, D. J. (2018). *Robot rights*. MIT Press.
Gunning, D. (2017). Explainable artificial intelligence (xai). *Defense Advanced Research Projects Agency (DARPA), nd Web, 2*.
Guzman, A. L. (2018). What is human-machine communication, anyway? In A. L. Guzman (Ed.), *Human-machine communication: Rethinking communication, technology, and ourselves* (pp. 1–28). Peter Lang.
Haake, M., Axelsson, A., Clausen-Bruun, M., & Gulz, A. (2015). Scaffolding mentalizing via a play-&-learn game for preschoolers. *Computers & Education, 90*, 13–23.
Heaven, D. (2019). Two minds are better than one. *New Scientist, 243*(3244), 38–41.
Humble, N., & Mozelius, P. (2019). *Teacher-supported AI or AI-supported teachers?* Paper presented at the European Conference on the Impact of Artificial Intelligence and Robotics 2019 (ECIAIR 2019). Oxford, UK.
Hwang, W.-Y., Chang, C.-B., & Chen, G.-J. (2004). The relationship of learning traits, motivation and performance-learning response dynamics. *Computers & Education, 42*(3), 267–287.
Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... Potapenko, A. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583–589.
Kaptelinin, V. (1992). *Integration of computer tools into the structure of human activity: implications for cognitive ergonomics.* Paper presented at the Human-computer interaction: Tasks and organisation. Proceedings of the sixth european conference on cognitive ergonomics.
Kaptelinin, V. (1996). Computer-mediated activity: Functional organs in social and developmental contexts. In B. A. Nardi (Ed.), *Context and consciousness: Activity theory and human-computer interaction* (pp. 45–68). MIT Press.
Kasparov, G. (2008). *How life imitates chess*: Random House.
Kasparov, G. (2019). AlphaZero and the knowledge revolution. In M. Sadler & N. Regan (Eds.), *Game Changer*: New in Chess.
Kellogg, D., & Veresov, N. (2019). Setting the stage. *LS Vygotsky's Pedological Works* (pp. 65–84). Springer.
Kennedy, M. L. (2019). What do artificial intelligence (AI) and ethics of AI mean in the context of research libraries? In E. A. Waraksa (Ed.), *Research Library Issues. RLI 299: Ethics of Artificial Intelligence.* Association of Research Libraries.
Kissinger, H. (2011). On China. New York: The Penguin Press.
Korf, R. E. (1997). Does deep blue use artificial intelligence? *ICGA Journal, 20*(4), 243–245.
Kurzweil, R. (1990). *The age of intelligent machines*: MIT press Cambridge.
Kurzweil, R. (2015). Superintelligence and singularity. In S. Schneider (Ed.), *Science Fiction and Philosophy: From Time Travel to Superintelligence* (pp. 146–170). Wiley-Blackwell.
Lazzeri, S. G., & Heller, R. (1996). An intelligent consultant system for chess. *Computers & Education, 27*(3–4), 181–196.
Lesgold, A. M. (2019). *Learning for the age of artificial intelligence: Eight education competences*: Routledge.
Levy, S. (2022). Blake Lemoine says Google's LaMDA AI faces 'Bigotry'. *Wired*. Retrieved from https://www.wired.com/story/blake-lemoine-google-lamda-ai-bigotry/. Accessed 9 July 2022.

Lindblom, J., & Ziemke, T. (2003). Social situatedness of natural and artificial intelligence: Vygotsky and beyond. *Adaptive Behavior, 11*(2), 79–96.
Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*: Pantheon.
Metz, C. (2016). The sadness and beauty of watching Google's AI play Go. *Wired*. Retrieved from https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/. Accessed 10 October 2020.
Molenaar, I., Van Boxtel, C. A., & Sleegers, P. J. (2010). The effects of scaffolding metacognitive activities in small groups. *Computers in Human Behavior, 26*(6), 1727–1738.
Mubin, O., Stevens, C. J., Shahid, S., Al Mahmud, A., & Dong, J.-J. (2013). A review of the applicability of robots in education. *Journal of Technology in Education and Learning, 1*(209–0015), 13.
Nagao, K. (2019). *Artificial intelligence accelerates human learning: Discussion data analytics*. Springer.
Nielsen, P. H. (2019). The exciting impact of a game changer: When Magnus met AlphaZero. *New in Chess, 8*, 2–10.
Nkambou, R., Bourdeau, J., & Psyché, V. (2010). Building intelligent tutoring systems: An overview. In *Advances in Intelligent Tutoring Systems* (pp. 361–375). Springer.
Norman, D. A. (1991). Cognitive artifacts. In J. M. Carroll (Ed.), *Designing interaction: Psychology at the human-computer interface* (pp. 17–38). Cambridge University Press.
Nussbaum, M., Alvarez, C., McFarlane, A., Gomez, F., Claro, S., & Radovic, D. (2009). Technology as small group face-to-face collaborative scaffolding. *Computers & Education, 52*(1), 147–153.
Nwana, H. S. (1990). Intelligent tutoring systems: An overview. *Artificial Intelligence Review, 4*(4), 251–277.
Okumura, Y., Kanakogi, Y., Kanda, T., Ishiguro, H., & Itakura, S. (2013). The power of human gaze on infant learning. *Cognition, 128*(2), 127–133.
Ormrod, J. E. (2016). *Human learning*: Pearson Higher Ed.
Reid, V. M., & Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *European Journal of Neuroscience, 21*(6), 1763–1766.
Rietveld, E., Denys, D., & Van Westen, M. (2018). Ecological-enactive cognition as engaging with a field of relevant affordances. In A. Newen, L. De Bruin, & S. Gallagher (Eds.), *The Oxford handbook of 4E cognition* (vol. 41–70). Oxford University Press.
Robbins, S. (2019). A misdirected principle with a catch: Explicability for AI. *Minds and Machines, 29*(4), 495–514.
Rolla, G., & Novaes, F. (2020). Ecological-enactive scientific cognition: Modeling and material engagement. *Phenomenology and the Cognitive Sciences*, 1–19. https://doi.org/10.1007/s11097-020-09713-y
Russell, S., & Norvig, P. (2014). *Artificial intelligence: A modern approach* (3rd ed.). Pearson.
Sadler, M., & Regan, N. (2019). Game changer. *New in Chess*.
Sætra, H. S. (2019). When nudge comes to shove: Liberty and nudging in the era of big data. *Technology in Society, 59*, 101130. https://doi.org/10.1016/j.techsoc.2019.04.006
Sætra, H. S. (2020). The parasitic nature of social AI: Sharing minds with the mindless. *Integrative Psychological and Behavioral Science, 54*, 308–326. https://doi.org/10.1007/s12124-020-09523-6
Sætra, H. S. (2021). Challenging the neo-anthropocentric relational approach to robot rights. *Frontiers in Robotics and AI*. https://doi.org/10.3389/frobt.2021.744426
Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., ... Graepel, T. (2020). Mastering atari, go, chess and shogi by planning with a learned model. *Nature, 588*, 604–609. https://doi.org/10.1038/s41586-020-03051-4
Shannon, C. E. (1950). XXII. Programming a computer for playing chess. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 41*(314), 256–275.
Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... Lanctot, M. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature, 529*(7587), 484.
Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., ... Graepel, T. (2017). Mastering chess and shogi by self-play with a general reinforcement learning algorithm. arXiv preprint arXiv:1712.01815
Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., ... Graepel, T. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science, 362*(6419), 1140–1144.
Strogatz, S. (2018). One giant step for a chess-playing machine. *New York Times*.
Strøm, O. K. (2019). PAPPA GIR Dataprogram ÆREN for magnus carlsens forvandling. Retrieved from https://www.vg.no/sport/i/naPnXL/pappa-gir-dataprogram-aeren-for-magnus-carlsens-forvandling. Accessed 10 October 2020.
Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*: MIT press.

Tomasello, M. (2009). *The cultural origins of human cognition*: Harvard university press.
van den Herik, H. J. (2018). Computer chess: From idea to DeepMind. *ICGA Journal, 40*(3), 160–176.
Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*: Harvard university press.
Vygotsky, L. S. (1987). *The collected works of LS Vygotsky: Problems of general psychology, including the volume thinking and speech*. New York: Plenum Press.
Vygotsky, L. S. (1989). *Concrete Human Psychology. Soviet Psychology, 27*(2), 53–77. https://doi.org/10.2753/RPO1061-0405270253
Vygotsky, L. S. (1997). *Educational psychology*. St. Lucie Press.
Vygotsky, L. S. (2019). The problem of the environment in pedology. *LS Vygotsky's Pedological Works* (pp. 65–84). Springer.
Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Transparent, explainable, and accountable AI for robotics.
Webb, K. (2019, November 27). A former world champion of the game Go says he's retiring because AI is so strong: 'Even if I become the No. 1, there is an entity that cannot be defeated'. Retrieved from https://www.businessinsider.com/deep-mind-alphago-ai-lee-sedol-south-korea-go-2019-11?r=US&IR=T. Accessed 10 October 2020.
Wertsch, J. V. (2002). Computer mediation, PBL, and dialogicality. *Distance Education, 23*(1), 105–108.
Wertsch, J. V. (2007). Mediation. In H. Daniels, Michael, Cole, & J. V. Wertsch (Eds.), *The Cambridge companion to Vygotsky* (pp. 178–192). Cambridge, Cambridge University Press.
Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100.
Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society, 20*(1), 118–136. https://doi.org/10.1080/1369118X.2016.1186713
Yun, G. (2017, May 27). Glorious still is Ke Jie despite 0–3 loss to AlphaGo. *China Global Television Network*. Retrieved from https://news.cgtn.com/news/3d49544f7a41444e/share_p.html. Accessed 10 October 2020.