Yoselin Chavez
TRA301
Bangalore
21 April 2024

<div align="center">TRA301 Final Paper</div>

There has been extensive use of AI in the field of education over the past decades, though development and implementation has had a noticeable massive uptake in the most recent past couple of years. With the open availability of services such as ChatGPT and Copilot, LLMs are easily accessible to the general public, regardless of any underlying understanding of what exactly they are interacting with. As such, students have begun using AI to help aid with their studies, from explaining concepts to writing entire essays, and teachers have been doing the same, using AI to help where they can from writing lesson plans to checking for plagiarism. And to a larger extent, AI is being implemented in the learning environment as a whole in official aspects. As it stands, there is little regulation on the usage of AI in the classroom, leaving discretion often to individual teachers, or to school administrators when there's more proactivity.

There has been some progress towards policies, such as the exploration of AI in education in terms of human rights, particularly in Europe, however the extensive debates as to what regulations should cover and look like causes actual implementation to be stunted. In spite of this, large scale rollout on the usage of AI in education has already started, such as experimental monitoring technology of the brain activity of students in China, and data gathering for Virtual Learning Environments meant to predict 'at-risk' students.[1] And aside from more organized implementation by say an entire school district or school board, individual teachers and students are also making the active choice to make use of AI when it comes to school work, particularly with online learning platforms such as Pearson. As a result, many questions are being raised on the matter, such as the ethical implications, the logistical implications, and the practical implications. Ultimately, the question boils down to whether AI can effectively be used in education in a way that is beneficial and sustainable for everyone in the long run.

The ways in which AI is being used for education in recent times are wide and varied. AI assisted apps, assisted simulations, chatbots, and e-proctoring are just a few examples of services under the AIED (AI in Education) umbrella of tools that are already being regularly made use of in both small and large scale cases.[2] Some teachers go out of their way to implement AI into their classrooms in small ways where there is no large scale roll out, and some students also use the publicly available services of their own volition to aid in completing assignments. This is particularly true for chatbots, which tend to be used for anything and everything, seen as all encompassing information banks by many of the general public. Meanwhile, larger scale implementations are becoming more common as well. In China, there was an AI system deployed that consisted of students wearing headbands with small electrodes meant to measure the electrical signals in students' brains. This allowed for measurement of concentration levels, with teachers being able to identify when students became distracted and

---

[1] Berendt et al., "AI in education", 313
[2] Holmes and Tuomi, "State of the art", 551

stopped 'paying attention'.[3] China in particular has had a lot of advancements in the idea of AI in education compared to other countries, and have been reportedly well received, such as the application of Personalized Learning Systems.[4] Similar rollouts can also be found elsewhere, such as in the UK, there is a system of data monitoring within a university's Virtual Learning Environment known as Moodle. This system uses data to detect the likelihood of an individual's ability to pass an assessment, or determine whether they are at risk of dropping out. It even dictates what a learner needs to study, or whether additional assistance is necessary.[5] Such extensive applications of AI technology have been becoming more common, and while some of the extreme technology is experimental at the moment, plenty is becoming permanent, such as the predictive models in the UK.

The effectiveness of these implementations are already being brought into question, however, as the claims of what AI can achieve in education become more ambitious and grandiose. As it stands, AIED is advertised for being an invaluable tool for educators, for a variety of purposes. Predictive models, for instance, are becoming increasingly popular as they can be used to track behavior of students to see which are at risk of abandoning their studies, or identify gifted students. Other lauded benefits include personalized learning materials, monitoring learning progress, assessments with feedback, and the ability to analyze scaled data for evaluation or administrative purposes, as well as automating repetitive tasks that take time away from teacher's interaction with students, such as marking, record keeping, and writing emails.[6] In greater capacities, AIED is being advertised in entirely revolutionary ways. Pearson, a large world wide digital learning platform, is just one of many proponents of this new technology, obviously with their own stake in the matter. Part of their argument for what AIED brings to the table is AIED for collaboration, which takes the idea that students learn better when learning together rather than individually, and using different approaches in order to create this inherently non-spontaneous learning environment for students.[7] This particular application of AIED has four main categories: Adaptive Group Formation, which uses knowledge about particular individuals to form groups best suited for particular tasks in an effort to group students of similar cognitive levels and interests or of complementary knowledge and skills as seen fit. Expert Facilitation, which monitors collaboration patterns for concepts students are struggling with as a group, as well as how much an individual is contributing to a group for grading purposes. Intelligent Virtual Agents are AI agents designed to mediate student interaction or contribute dialogues as say an "expert", or someone for the participants to teach the material to. And finally, there's Intelligent Moderation, which summarizes discussions in large groups for a human to guide the students, with alerts for incidents such as widespread misconceptions or excessive off topic discussions.All of these methods are meant to take the collaborative, social benefits of human interaction and integrate them with AI to create an optimal learning environment for students that is practical in implementation for teachers with their limited man power. AIED research has found that "expert systems" in particular can in fact prove effective, with students interacting with such systems consistently outperforming other groups without the

---

[3] Berendt et al., "AI in education", 316
[4] Zhang and Aslan, "AI Technologies", 6
[5] Berendt et al., "AI in education", 313
[6] Zhang and Aslan, "AI Technologies", 6
[7] Luckin and Holmes, "Intelligence Unleashed", 26-27

supplemental support in experimental research, at least when it comes to mathematics learning achievement.[8] The claims suggest that not only can AI mimic the current favored learning environments for students without the same amount of labor necessary, but could even be an improvement in some ways. Pearson continues with the claims of what AIED can achieve, such as with immersive environments in virtual realities to help low-achieving students build confidence by allowing them to, say, become successful virtual scientists, or help students navigate the world of bullying with effective coping strategies in simulations, all of which aren't possible without this advanced technology.

There are also the slightly more economical benefits touted by AIED programs. Teacher burnout is considered a leading cause of professionals currently leaving the field, due to stress and workload. This, along with the often unsatisfactory wages relative, have resulted in teacher shortages, with "33 countries [that] do not currently, and will not have, enough teachers to provide every child with a primary education by 2030".[9] To certain degrees, AI is even advertised as having the potential to serve as personal tutors, bridging the gap for those unable to afford traditional private tutors for children and making personalized education more accessible to all students. It's also supposed to eliminate the bias and discrimination against those of minorities and lower socioeconomic classes with the objective, unbiased position of machine models.

Despite all these claims, however, the actual feasibility of these capabilities is another matter. To start, there are questions as to the ethical implications over the growing use of AI not just in the classroom environment, but in general. In terms of human rights, the issues of "Right to human dignity, Right to autonomy, Right to be heard, Right not to suffer from discrimination, Right to data privacy and data protection, [and] Right to transparency and explainability" are all brought up as concerns.[10] These ideas of basic rights bring into question how AI should be implemented, whether it should be implemented at all, and how data in particular should be handled.

With the lack of regulations comes the lack of definition on ownership when it comes to data, and the information that is allowed to be collected. For instance, in the case of the Moodle prediction system, there is concern that the systems that are meant to support students in passing assessments and measuring their progress can be easily manipulated for any number of reasons. In the cases where a school district gets more funding with higher test scores, they might manipulate the system to artificially inflate scores. Conversely, if improvement is rewarded, there are incentives to downplay initial performances, and raise later ones.[11] There's also concerns that schools could use predictive modeling to prioritize "teaching to test", particularly with tutors and cram schools that have the goal of helping students get into particular schools. There's already evidence of such practice with schools in more affluent areas of the UK, where teachers have been teaching specifically for university entrance exams. Traditional assessments have already been under fire for not being completely effective means of measuring learning outcomes, such as with the SATs, and are considered to even contribute to existing systemic discrimination in the education system. Further, while there are currently

---

[8] Zhang and Aslan, "AI Technologies", 5
[9] Lucking and Holmes, "Intelligence Unleashed", 44
[10] Holmes and Tuomi, "State of the art", 558
[11] Berendt et al., "AI in education", 315

requirements to acquire consent from individuals before gathering and storing information about them, it's not entirely effective in making sure the data is used properly. The agreements and waivers people sign to allow the collection and use of their data can have extremely vague wording, resulting in companies essentially being able to do whatever they want with the information once they have it. Lifelong data is a particular issue, especially since many current and developing models boast predictions about an individual's future. Thus, there's the issue of not giving people the option to have details forgotten about them, and being constrained to facts and a past they might not necessarily want to be associated with anymore. In education in particular, there's special focus on childhood and adolescence despite the fact that growth doesn't necessarily happen for everyone at that stage in life. That could lead to what people consider to be unequal or inaccurate profiles about themselves that they are unable to escape from. This constant gathering of information could also eliminate the "safe space" aspect for many children, with everything being recorded and stored as data, making people less likely to open up in these kinds of environments as they otherwise might.

While "opting out" might seem like the solution to a lot of these issues, that isn't necessarily an easy option for many. For those under the legal age, those decisions have to be made by their parents or guardians, leaving many children forced in or out of programs without much consideration for their position. Further, there's often strong social pressure not to opt out, leaving many parents to consent simply to bend to societal norms.[12] It's a paradox of human rights and obligations, since in most every developed country at the very least, children are forced to attend school where they need to participate in learning activities, assessments, and more. There's already debate as to the ethics of forcing these things upon people, and the issue is only exacerbated by the presence of constant data gathering. The degree to which people can opt out is becoming more limited, with the difficulty and effort required to do so is also increasing. Once again, there's also the economical aspects to consider. With the lack of regulations on AI and the data that comes with it, the practice of treating personal data as property to buy and sell is well within the law for the most part. This allows people to sell their data and data rights, which has already been occurring in daily life, and is of course an option that is likely to be more often chosen by the poorer and less educated population.

People have also expressed concerns over the ubiquitous existence of tracking, as well as the ever growing dependency on AI. There's already concerns about companies tracking and selling personal data from internet use and the like, but the same can also occur for data meant to be used for education models. Roll outs in China have included data chips embedded into children's uniforms for facial recognition and locational data, analyzing engagement and interaction. While these extremes have not yet become commonplace, that eventually is becoming more and more possible. And once again, the option to opt out of such programs is becoming more difficult for the average citizen, with lots of incentivization not to opt out at all.

The effects of long term interaction with AI is also part of the debate. In particular, there is the question of whether the extensive use of AI will ultimately lead to long term cognitive decline, such as Socrates' famous argument that writing led to forgetfulness, and the fact that studies have suggested the use of GPS leads to decline in spatial memory.[13] There have also been concerns that over reliance on AI will result in overall memory deterioration due to

[12] Berendt et al., "AI in education", 316
[13] Holmes and Tuomi, 'State of the art", 560

underutilization, and impaired critical thinking due to the overuse of AI preventing the development of critical thinking skills.[14] The concerns are valid, considering the all true nature of developmental stagnation due to underuse, and is a current leading argument against the use of AI in the classroom. Despite this, however, there have been strong counter arguments, such as the idea that rather than memory simply deteriorating as a whole, there's evidence that human brains are simply adapting to the information that is easily available through services such as the internet, and instead prioritizing information that needs to be retained through proper memory.[15] There's also the evidence that things such as smartphone alerts allow cognitive resources for other tasks. There have also been suggestions that a proper integration of blended learning approaches, with the incorporation of critical thinking and problem solving abilities in conjunction with services such as ChatGPT, can mitigate negative effects.[16]

Looking at the evidence through these facts another way, rather than simply becoming worse, the human brain is evolving and adapting to the tools available, and choosing what to prioritize. With more information and capabilities becoming easily available, the human brain can shift focus elsewhere, and perform different tasks. After all, even if a tool can do something, it must still be utilized properly. People need to develop the ability to actually make proper use of the tools and to develop them further, so instead of focusing on completing the same tasks these readily available tools can accomplish, the human brain can focus instead on adapting to use the tools most effectively, as well as finding new ways to make use of it. That is how humans progress, after all, and so of course the supposed norm of what a person can and should be able to do is bound to change with the times as well as the technology available.

There's also the ultimate psychological impacts of this growing use of AIED, and really AI usage in general. On the proponent of AIED side, there's many inherent positive impacts to be had from the growing use of AI in the classroom, such as research that has found increased use of AIED tools has resulted in reducing anxiety for students, particularly with the "expert system" with math, and Teachable Agents designed to act as private tutors in science.[17] These tools are also advertised as giving students the security of round-the-clock support and supposed accommodative potential for those in need, reportedly resulting in reduced pressure for both students and teachers.[18] And again, the supplementation capabilities of AIED are considered a potential solution for the burnout and overwork many teachers are currently experiencing, which should result in better overall experiences for both educators and students. Despite these supposed benefits, however, there are still concerns. While there have been developments to implement social and collaborative aspects to AIED, many of the current ways in which these tools are being used effectively involve learning methods that are considered isolating. Such as, most students using ChatGPT alone without supervision, rather than those collaborative specific AIED programs touted by Pearson. Thus, some people argue that while these tools might be helpful for things such as homework, they perhaps might not be best suited for classrooms, which are inherently supposed to be more social places.[19] Anecdotally, many teachers have also spoken out as to how while AI might help save time with its capabilities, it

---

[14] Bai et al., "ChatGPT", 2
[15] Sparrow et al., "Google Effects on Memory", 778
[16] Bai et al., "ChatGPT", 6
[17] Zhang and Aslan, "AI Technologies", 5
[18] Bai et al., "ChatGPT", 5
[19] Holmes and Tuomi, "State of the art", 559

can feel distant, cold and impersonal, particularly for cases that involve direct parent-teacher correspondence.[20] Many teachers who value the interpersonal connections with their students have opted to minimize their usage of AI, despite the extent to which it could lower their workload. Alternatively, this means other teachers might, for whatever reason, decide to make full use of AI at the sacrifice of what is often considered an all important aspect of the education process while AI is seemingly unable to make up for that emotional feature of education. And while those rollouts of experimental technology in China have resulted in higher scores, stress levels have not been measured.[21] The extensive supervision could undoubtedly be placing strenuous amounts of stress on students, but could also be impacting teachers considering how this information could be used to evaluate the performance of teachers and their abilities in the classroom. And once again, the culmination of concerns people have over AI can remove that safe space aspect of the education system, ultimately raising stress levels and removing the casual social aspects for students regardless. After all, the purpose of education isn't simply to set students up for getting a job, but also exploring passions, finding satisfaction in the work they accomplish, and becoming a well rounded, moral person.

All of the debate as to whether we should use AI in the classroom and to the extent of it is also assuming that these tools are actually capable of the tasks they would be used for. As it currently stands, "general" AI does not exist, which essentially amounts to the concern of AI becoming too intelligent and leading to the "singularity" of AI replacing humans. In fact, there is considered to be a rather limited understanding of intelligence itself, which tends to get glossed over in the AI community. That limited understanding is at odds with the need to understand the limitations of AI's capability to simulate ranges of intelligences, as AI tends to follow information processing models that consider intelligence to be purely brain based, readily available for transfer to software. Studies have shown, however, that intelligence isn't such a simple concept, and instead consists of common sense, emotions, and irrational thoughts that are intrinsically interwoven with the complexities in individual lives.[22] Most AI as is are thus domain specific rather than all encompassing, sentient creations, and depend on both knowledge of the world as well as algorithms to process that knowledge.[23] In fact, there might be greater concern over people thinking AI is sentient when it is not, rather than the actual issue of AI becoming sentient, particularly as people become more reliant on it for their day to day lives.[24] After all, while the outputs from services such as ChatGPT might provide convincing looking results, they aren't necessarily accurate, consistent, or reliable. This issue has been labeled "AI Hallucation", where AI produces content that lacks basis or accurate data to corroborate it, often even creating false statements and evidence from nowhere.[25]

For instance, when it comes to math, while ChatGPT-4 has been showing the most consistent and correct results amongst the most popular of chatbots, success rates are still varied. An experiment performed by engineering researchers tested the math capabilities in particular of three of the most popular chatbots currently available: ChatGPT-3.5, ChatGPT-4,

---

[20] Will, "With ChatGPT"
[21] Berendt et al., "AI in education", 316
[22] Selwyn, "Future of AI", 623
[23] Luckin and Holmes, "Intelligence Unleashed", 15
[24] Selwyn, "Future of AI", 622
[25] Plevris et al., "Chatbots Put to the Test", 967

and Google Bard.[26] The methodology of the experiments involved sets of math and logic problems with one set having answers available online, and another completely unique, with no results available on the internet. Each chatbot was presented with each question three times in order to test consistency and accuracy, with the responses scored based on correctness. There were rather varied success rates across the board, with ChatGPT-4 generally performing better than ChatGPT-3.5 in all circumstances. Meanwhile, Bard generally did worse than both in general, only having the advantage with problems easily accessible through the internet and being able to pull the correct answers from there. Consistency issues were also present, as even ChatGPT-4 will often get the correct answer on say 1 out of 3 tries, getting something wrong the first time, right the second, then incorrect again the third. While the three popular chatbots (GPT 4, GPT 3.5, and Bard) seemed to mostly be capable of solving and understanding purely logical problems, the results for math specific problems were less than preferable. For instance, with a set of math questions not available on the internet, success rates consisted of 80% for ChatGPT-4, 53.3% for ChatGPT-3.5, and 40% for Google Bard. For the math questions with solutions on the internet, the rates consisted of 80% for Google Bard, 62.2% for ChatGPT-4, and 42.2% for ChatGPT-3.5. These results show that while there can be a significant amount of success from using these AI models for common homework, it's quite dependent on the subject and availability of the material. In short, a person must be aware of which model to use for their specific purposes in order to get the most reliable answers. And no matter what, these chatbots output their results with full confidence and "evidence", leaving users to have to fish out the accurate information from the inaccurate. Despite these facts, it still stands that many students in particular use these services for help with math homework, often completely unaware as to the intricacies involved with achieving proper results. These convincing but inaccurate answers would simply impede progress and development of students rather than supporting them. There's also predictive models that have been pushed more as of late, despite the dubious validity of their results. A Princeton University study challenged teams of data scientists, statisticians, AI and machine learning researchers to predict various life outcomes for children, such as grade point averages and their perseverance with schoolwork. Despite 13,000 data points on over 4000 families stretching back over fifteen years, every team failed to develop models that were even moderately successful.[27]

  This algorithmic and information based aspect of machine learning models also means that these systems are dependent on history, as well as their programming. AI products are bounded mathematical systems, and there's the fundamental question of the extent to which education and everything that comes with it can be properly quantified and represented in data, as well as the parts that might need to stay human dependent. Even if AI can reach a point where it can effectively answer problems for students and do what it is advertised to do, due to the very nature of what AI is, 'innovation' in the fields it is extensively implemented in can stagnate. Most AIED tools as they currently exist tend to drive homogenisation of students, meaning they provide minimal actual personalization, and aim primarily to fit students in the correct boxes to pass their exams. This is of course due to the fact that most AIED system

---

[26] Plevris et al., "Chatbots Put to the Test"
[27] Selwyn, "Future of AI", 623

individual pathways are based on averages and data from proper learners, meaning evolution in the thinking for how students should be evaluated could significantly decline.[28]

And of course, the reliance on past data brings into question whether AI is actually inherently unbiased, or if it'll simply perpetuate existing inequalities and discrimination. Considering the machine learning models require large amounts of data to be trained, the majority of models are trained in English, with information usually gathered from the United States using studies performed primarily in smaller scales, making the developments not inherently transferable to other countries that have different learning systems.[29] Even when data is gathered from other countries, it doesn't inherently make all models more universal as they tend to be very area and locationaly focused, continuing to make this data not universally applicable. Aside from this, even within individual countries with larger data sets, the data gathered and used is usually fundamentally flawed. There has been evidence of algorithmic discrimination, such as non-native accents being judged for cheating on tests, computing higher grades for students whose profiles have historically been more likely to receive high grades (such as based on race, socioeconomic status, etc.), facial recognition software that is unable to properly recognize students of color, and eye-tracking data that is based on steady gazes as well as the behaviors associated with neuro-typicals.[30] Plus, with the opting out behaviors, there's also issues of biased over and under representation of specific groups based on tendencies and likelihood of opting out from data collection. The issue is particularly troublesome considering the difficulty in understanding exactly which groups are being misrepresented, as well as in the ways they are being misrepresented, along with the fact that this occurrence might not even be garnering attention in many parts of the machine learning community.[31] The result is engineered and invisible inequality, in which the ways students are being discriminated against are magnified and automated by the software meant to help them succeed.

As it stands, while AI has plenty of potential to aid students and teachers in the education system, and there is even evidence of those benefits becoming apparent, there is still a ways to go before AI can even deliver the results that it has promised thus far, particularly in fair and balanced ways. In order to reach that point, it's vital to address the systemic biases present in the data used to train these models, as well as the potential abuse of these systems by those working with them behind the scenes. And even aside from these issues, the ethical implications of the handling of both data and this software itself are still constantly being debated. There's particular concern as to the extent AI should be implemented within the education sector, aside from the practicality aspects. While AI is convincing, it has proven to be limited in its actual consistency and accuracy, as well as its domain specific nature. This leads to the conclusion that different machine learning models need to be utilized in the appropriate circumstances, and as such, a fundamental understanding of what each model is capable of is necessary. This understanding is often lost to the general public, such as students and teachers making use of models individually, leading to more detrimental results rather than benefits. And while there are vast claims as to the applications of AIED for student learning, the current state

---

[28] Selwyn, "Future of AI", 628
[29] Holmes and Tuomi, "State of the art", 560
[30] Selwyn, "Future of AI", 624
[31] Berendt et al., "AI in education", 314

of things is such that machines cannot entirely replace human teachers for a lot of vital aspects, such as the social and emotional aspects. Even if they could, it does not mean that they should. After all, making machines that are no worse than humans isn't justification for completely adopting AI into the educational system, considering how humans themselves are still wrought with their biases and senselessness. The conclusion is that ultimately, the degree to which AI can and should be used in education will need to be discussed and monitored, dealt with on a case by case basis. There is no clear cut "machines bad, humans good" or vice versa, and instead, a hybrid, integrated system will ultimately be what works best, requiring actual understandings of what these models are capable of and how to use them.

Works Cited

Bai L, Liu X, Su J. ChatGPT: the cognitive effects on learning and memory. *Brain-X*. 2023;
      1:e30. https://doi.org/10.1002/brx2.30

Berendt, B., Littlejohn, A., & Blakemore, M. (2020). AI in education: learner choice and
      fundamental rights. *Learning, Media and Technology*, 45(3), 312–324.
      https://doi.org/10.1080/17439884.2020.1786399

Betsy Sparrow et al. ,Google Effects on Memory: Cognitive Consequences of Having
      Information at Our Fingertips. *Science* 333,776-778(2011).DOI:10.1126/science.1207745

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European
      Journal of Education*, 57, 542–570. https://doi.org/10.1111/ejed.12533

Luckin, R., & Holmes, W. (2016, February 22). Intelligence Unleashed: An argument for AI in
      Education. *UCL Discovery*. https://discovery.ucl.ac.uk/id/eprint/1475756/

Plevris V, Papazafeiropoulos G, Jiménez Rios A. Chatbots Put to the Test in Math and Logic
      Problems: A Comparison and Assessment of ChatGPT-3.5, ChatGPT-4, and Google
      Bard. *AI*. 2023; 4(4):949-969. https://doi.org/10.3390/ai4040048

Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of
      Education*, 57, 620–631. https://doi.org/10.1111/ejed.12532

Will, M. (2023, March 1). With CHATGPT, teachers can plan lessons, write emails, and more.
      what's the catch? *Education Week*.
      https://www.edweek.org/technology/with-chatgpt-teachers-can-plan-lessons-write-emails
      -and-more-whats-the-catch/2023/01

Zhang, K., & Aslan, A. B. (2021, June 9). AI Technologies for Education: Recent research &
      future directions. *Science Direct*.
      https://www.sciencedirect.com/science/article/pii/S2666920X21000199