Journal of Learning and Teaching in Digital Age, 2021, **6**(2), 97-100
https://dergipark.org.tr/en/pub/joltida
ISSN: 2458-8350 (online)

Perspective

Do Less Teaching, Do More Coaching: Toward Critical Thinking for Ethical Applications of Artificial Intelligence

Claire Su-Yeon Park[a], Haejoong Kim[b], Sangmin Lee *[c]

[a](ORCID ID: 0000-0002-2109-9885), University of Alberta Faculty of Nursing, Edmonton, Canada, suyeon4@ualberta.ca
[b](ORCID ID: 0000-0002-3727-094X), Material Handling Automation Group, Samsung Electronics Co., Ltd., Hwaseong-si, Gyeonggi-do, Republic of Korea, haejoong.kim@gmail.com
[c](ORCID ID: 0000-0002-5215-2546), School of Information Convergence, Kwangwoon University College of Software and Convergence, Seoul, Republic of Korea, smlee5679@gmail.com
*Corresponding author

ARTICLE INFO

Received: 27 September 2020
Revised: 03 March 2021
Accepted: 13 March 2021

Keywords:
Artificial intelligence
Critical thinking
Intellectual humility
Intellectual courage
Coaching

ABSTRACT

There have been discussions suggesting an ethics committee be established which would oversee humanity's efforts in Artificial Intelligence (AI) and its applications to our society. This concern arises mostly because of the limitations of existing data used for the development of AI algorithms that intrinsically reflect unfair and discriminatory factors of the real world in which we live. However, it is hard to find a paper that philosophically addresses a pedagogical issue about the necessary shift from strict teaching to informative guidance: i.e., a conversation about developing the discernment and critical thinking skills which would allow people who use AI-integrated services to themselves monitor the AI's ethical applications and thus secure the well-being of our society. This paper is differentiated from other papers in that it sheds light on the social problems that can arise if people become uncritically compliant with unethical and indiscriminate applications of AI, and it conveys the lesson that contemporary ordinary citizens should be alert to these pitfalls as well.

INTRODUCTION

As social distancing has become a survival condition in the era of new coronavirus infections, "classroom-shared education" with emphasis on collective efficiency has gone, and "remote universities" (equipped with Artificial Intelligence (AI) technology and thoroughly specialized individuals, while also being inexpensive and not limited by time and space) are rapidly driving the future of education (Kim, 2020). Taking it one step further, Virtual Reality (VR)-based education platforms which incorporate the real economy are being actively researched and developed (Park, 2018). Augmented Reality (AR)-based education concepts are also emerging as a hot ticket to enhance the effectiveness and attractiveness of both teaching and education (Horváth, 2018). Thus, it may be possible for educators and students to perform experiments and practice performances in virtual space in the near future. This is very encouraging to theory-driven practical scholarship such as nursing science because practicum is an indispensable condition of these disciplines. Such advanced technologies-integrated, student-centered remote education systems may accordingly force a crisis in brick-and-mortar institutions of higher education. It is not an exaggeration to say that in a virtual space, unlike the current education system, infinite competition takes place within (and outside) these virtual education systems in order to attract as many students as possible--because they can meet their educational needs beyond the constraints of time and space. That crisis is not so far in the future.

We seem to be in a time of transition. Students' anger and protests for tuition refunds are a reaction to these slapdash and poorly prepared, low-quality online education triggered by the coronavirus COVID-19 pandemic—a signal of the oncoming crisis in higher education. We, as an interdisciplinary team, so-called "SECURE Team For You" (SweEt spot ConsUlting REsearch Team For the next generation, You) to develop Park's (2017) Optimized Nurse Staffing (Sweet Spot) Theory-driven AI decision-making support system, have had in-depth discussions about the meaning and role of future higher education educators in order to be proactive, not reactive, to survive the forecasted crisis. This perspective article is a report summarizing our iterative discussions from July 2020 to February 2021, which aims to provide helpful insight into the future higher education educators' preparation of pedagogics, written based on our interdisciplinary consilience -driven research experiences.

DISCUSSIONS

New paradigm in pedagogy is emerging: From teaching within the existing knowledge transfer to coaching/guiding a student-led critical thinking

We believe that these changes in social phenomena suggest that the main purpose of education needs to be shifted from teaching within the existing knowledge transfer to coaching/guiding a student-led critical thinking. Following a thoroughly personalized AI recommendation system, so-called "Google's Deep Neural Networks for YouTube Recommendations" (Covington, Adams, & Sargin, 2016), the rising generation is already moving to find platforms and content that suit their tastes and meet their individualized needs. They are already employing self-directed learning, breaking away from the traditional mold of students who remain in passive roles.

Risks posed by uncritical conformity of AI algorithms may threaten the well-being of society.

The problem lies in not seeing the forest for the trees. There are too many virtual education systems, and we are picking it indiscriminately. Attracted by automated algorithms someone else has devised, Lacan's Desire Theory's philosophical proposition, "Humans desire the desire of others" (Belsey, 1993), may now cause students to desire the desires of AI algorithms, more specifically, the desires of developers or business owners which fabricate such algorithms. As such, if students just follow AI systems, which have a certain propensity in forecasting or decision-making and lead to bias, without any critical thinking, this mindless following may consequently cause serious problems to the well-being of our society.

Imagine developers or hackers with nefarious agendas manipulating or causing malfunctions in existing AI systems for the sole purpose of creating specific bias (Madry, Makelov, Schmidt, Tsipras, & Vladu, 2017). What a horrible future we have to look forward to!

Meanwhile, Facebook's AI chatbot "Bob and Alice" was terminated because of an inappropriate use of its own language (Mamiit, 2017). Despite being trained to negotiate with humans using plain English, the AI used weird English words that, in context, could not be understood by native English-speakers (Mamiit, 2017). The case made ordinary citizens as well as the Facebook's AI developers feel embarrassed and even frightened by the possibility of one day losing control over decision-making in their daily lives due to the 'lost in translation' effect of human-AI interaction (Mamiit, 2017). Humans could not understand why AI systems yielded the result output, nor could they intuit these hidden information processes, also called the 'black box of deep learning.' (Mamiit, 2017; Yu, 2019). While these limitations are currently being overcome by rapid advances in eXplainable AI (XAI) technologies, there is still a long way to go (Yu, 2019). And, of course, if the AI agents were permitted to speak in their own programming language, humans would not be able to follow what the AI agents talk about because humans can't understand their 'words' nor the context of the conversation (Mamiit, 2017). Accordingly, humans would then be excluded from the decision-making process; from the standpoint of the AI agents, human intervention would be unnecessary and even inefficient. Thus, human beings would lose control over their own lives, uninformed and in the dark.

South Korea's AI chatbot "Lee Luda (Luda)" was also discarded because of its discriminatory and vulgar statements as well as privacy breach allegations (Figure 1). This happened mostly because of the limitations of existing data that intrinsically reflect the unfair and discriminatory factors already existent in the real-world culture. The Korea Personal Information Protection Commission (KPIPC) eventually slapped a total fine of KRW 13.3 million on the Scatter Lab, the developer of AI chatbot "Lee Luda" on April 28, 2021 (KPIPIC, 2021).

Both of these AI examples were strongly criticized for threatening social safety as well as the common good and were consequently pulled from our society. Both of them also proves how important humanity's critical thinking is to the realization of social justice, even in the age of AI. Computers could not do the work for us; they may merely reflect our own biases.

Figure 1. Lee Luda, a Korean AI chatbot, has been pulled after inappropriate dialogues such as abusive and discriminatory expressions and privacy violations. This case is thought-provoking in that it has raised moral concerns about data integrity for AI's ethical applications. Images were retrieved from the Scatter Lab (https://luda.ai/) and The Korea Times (http://m.koreatimes.co.kr/pages/article.amp.asp?newsIdx=302390) on 28 Feb 2021, and the first author translated text messages from Korean into English.

Virtual education may make differentiating between students' own desires and the desires of others impossible.

School allows students to differentiate between their own desires and the desires of others. It is possible to provide a holistic education which develops students who examine themselves and their motivations before taking action, students who shift from external rewards to internal. The sense of belonging together (i.e. solidarity), including social and community consciousness, cultivated by resolving conflicts in the relationship of religious leaders or accountability partners can be achieved along with knowledge transfer within the school system. However, with the advanced technologies-integrated, student-centered remote education system, such holistic education is virtually impossible. Thus obscured by convenience, we may be losing our autonomy and humanity and even self-esteem.

Particularly, uncritically relying on AI technologies in healthcare may cost us dearly.

The healthcare science community should particularly be more attentive to such conveniences; there is no shortcut without critical thinking. Uncritically relying on AI technologies in healthcare may cost us dearly: irreversible hurt (such as the death of a loved one) or damages that can't be easily recouped (such as physical/social disability, loss of job, and/or deteriorated quality of life), not to mention the possible psychological discomforts. We must keep in mind that there can exist no perfect solution. No matter how skilled our physicians or how elaborately developed our AI systems, nothing can eliminate many uncontrollable and unpredictable variables that intervene in any given situation. Thus, as a gatekeeper at the forefront of the healthcare delivery system, healthcare professionals should be prepared to "criticize" AI-driven solutions by detecting and preventing errors or biases because automated decision-making support systems will be applied to healthcare settings in omnidirectional ways soon if not now. Such preparedness will lead to securing patient safety and ensuring patients' well-being from possible disasters caused by human or AI errors, or both.

AI should not be a substitute for humans, but a complement in the form of *Intelligence Amplification*.

For the same reason, AI needs to be highlighted instead as Intelligence Amplification (IA), especially in the healthcare field. AI improves upon the acquisition of knowledge that is based on existing data. However, the AI's dependency on the existing data makes it impossible to generate a new knowledge set or a theory that did not exist before (Kong, 2018). Thus, academic and practical advances of AI may hinder the cultivation of new knowledge, a cultivation which requires human creative intervention. It consequently suggests that an AI's recommendations should not be viewed as the only right or optimal solution to the critical healthcare issues; also, the AI may need to be updated as technology advances. Richard Phillips Feynman, a theoretical physicist, says that we cannot create without understanding (Feynman, 2018). Even humans have not fully understood the world yet, so could an AI (coded by humans) truly produce anything new or creative? In addition, AI cannot understand the contexts of the many critical healthcare issues. Humans can understand and recognize that even the same expressions have different meanings depending on the given context and further separate out each meaning. In some cases, emotions can be mixed, confused, or contradictory, and only human beings can recognize and disentangle them. Unique human abilities (such as creativity, cognition, consciousness, analogical reasoning, or intuition) are something that (current) AI technologies can never express. In short not all causal chains in the world can be reduced to logic (Gödel, 1931). Thus, rather than leaving all the decision-making to an AI, healthcare professionals need to vigilantly monitor AI so that it can be a decision support system that enables people to make the most optimal decisions about critical healthcare issues with the least amount of human error as possible. AI should not be a substitute for humans but a complement as Intelligence Amplification (IA), particularly in healthcare.

Contemporary ordinary citizens should be alert to unethical and indiscriminate applications of AI as well.

All of the above remind us of the importance of an intellectual curiosity that motivates the pursuit of new knowledge; the intellectual humility that tolerates the idea that our thoughts and judgments may be wrong, accept others' opinions, and consider their criticism; the intellectual courage that frees us from the hegemony-based system in both academia and practice and challenges the existing and even well-established scientific methods and findings; and the developmental maturity that lets go of long-held beliefs (Jones et al., 2019; Park, 2018). This is because these qualities are a prerequisite for critical thinking. Above all, rapid advances in AI technologies create active, contemporaneous, continuous innovations in all disciplines which incorporate computer science. This supports why we place great emphasis on the contemporary importance of coaching student-led critical thinking rather than teaching within the existing knowledge transfer as virtues and values we should pass on to the next generation. It is no exaggeration to say that ordinary citizens should be alert to unethical and indiscriminate applications of AI as well considering that the speed of research and development of AI technology is very fast right now.

CONCLUSION

In an era in which problem-solving and design-thinking skills are becoming more and more important, we, educators, need to do more coaching and less teaching to help our students develop the necessary discernment and critical thinking skills. These advanced skills will in turn be reflected in AI systems, the ethical applications of which ensure the well-being and security of our society. We also need to encourage students to blaze their own trail and venture out in search of new knowledge and solutions rather than teaching old tricks. We might very well solve some long-standing global conundrums (say, the nursing shortages) in the world's healthcare systems by steering the students towards the intellectual curiosity, the intellectual humility, the intellectual courage, and the developmental maturity to achieve something new.

Authorship Contributions: We declare that all authors have contributed to this study and meet the criteria for the ICMJE Authorship guidelines on authorship. All authors have approved the final article and it has been confirmed that all authors have been listed as authors.

Successful approval of Institutional Review Board: This article is a perspective paper and did not utilize any data related to human subjects.

Funding Statement: The present research has been conducted by the research grant (no specific grant numbers) of Kwangwoon University in 2020. The recipient was Dr. Sangmin Lee. The funder (https://www.kw.ac.kr/en/index.jsp) did not and will not have a role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interest Statement: No conflict of interest has also been declared by the authors.

REFERENCES

Belsey, C. (1993). Desire in theory: Freud, Lacan, Derrida. *Textual Practice, 7*(3), 384-411, doi:10.1080/09502369308582173

Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for YouTube recommendations. *Proceedings of the 10th ACM (Association for Computing Machinery) Conference on Recommender Systems (RecSys '16)*, New York, NY, USA, 191–198. doi:10.1145/2959100.2959190

Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, *38*, 173-98. doi:10.1007/BF01700692

Feynman, R. P. (2018). *Classic Feynman: All the adventures of a curious character.* Edited by R. Leighton, and Translated by H. B. Kim and S. W. Hong. Science Books.

Horváth, I. (2018). Evolution of teaching roles and tasks in VR/AR-based education. *Proceedings of the 2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, *Budapest, Hungary,* 000355-000360. doi:10.1109/CogInfoCom.2018.8639907

Jones, T., Willis, E., Lopes, M. A., & Drach-Zahavy, A. On behalf of the RANCARE Consortium COST – CA 15208* (2019). Advancing the science of unfinished nursing care: Exploring the benefits of cross-disciplinary knowledge exchange, knowledge integration and transdisciplinarity. *Journal of Advanced Nursing,* 75, 905–917. doi:10.1111/jan.13948

Kim, J. H. (2020, April 15). [Kim Jung-Ho's Strategy in the AI Era] If you choose a future university that will come soon… Artificial Intelligence YouTube University! *ChosunMedia.* Retrieved from https://www.chosun.com/site/data/html_dir/2020/04/14/2020041403594.html (assessed on 13 Sep 2020)

Kong, B. (2018). *A dictionary of common sense in the fourth industrial revolution*. Gilbut.

The Korea Personal Information Protection Commission (KPIPC). Disposal of sanctions such as fines and fines to Scatter Lab, a developer of 'Lee Luda,' by the Korea Personal Information Commission. Available at: https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=7298 (accessed April 28, 2021).

Madry, A., Makelov, A., Schmidt, L., Tsipras, D., & Vladu, A. (2017). Towards deep learning models resistant to adversarial attacks. *arXiv preprint arXiv:1706.06083*

Mamiit, A. (2017, July 30). Facebook AI invents language that humans can't understand: System shut down before it evolves into Skynet. Retrieved from https://www.techtimes.com/articles/212124/20170730/facebook-ai-invents-language-that-humans-cant-understand-system-shut-down-before-it-evolves-into-skynet.htm (assessed on 03 Mar 2021)

Park, C. S. (2016). Nursing education's future: Blended learning. *Journal of Learning and Teaching in Digital Age, 1*(1), 2. Retrieved from https://dergipark.org.tr/en/pub/joltida/issue/55462/760051 (assessed on 03 Mar 2021)

Park, C. S. (2017). Optimizing staffing, quality and cost in home healthcare nursing: Theory synthesis. *Journal of Advanced Nursing*, *73*(8), 1838-1847. doi:10.1111/jan.13284

Park, C. S. (2018). The Dark shadow of Virtual Reality. *Journal of Learning and Teaching in Digital Age, 3*(1), 1-2. Retrieved from https://dergipark.org.tr/en/pub/joltida/issue/55469/760089 (assessed on 20 Sep 2020)

Park, C. S. (2018). Thinking 'Outside the Box' *Journal of Advanced Nursing*, *74*(2), 237-238. doi:10.1111/jan.13312

Yu, J. Y. (2019, April). Data science in practice: 'You need to put AI in your organization's decision making.' *Dong-A Business Review*, *1*(270). Retrieved from https://dbr.donga.com/article/view/1201/article_no/9086