Human-AI Decision-Making in Higher Education

Robert Houghton, Xinrui Zhai, Zhuojun Li & David R. Large

Human Factors Research Group, University of Nottingham, UK

SUMMARY

Findings from an online study and focus groups show that students and staff were receptive to AI decision-making in higher education, highlighting transparency and equitability. Even so, students expressed a desire for creative work to be seen by a 'fellow human', suggesting potential for a collaborative human/AI approach. However, a follow-up study revealed that this solution was in fact perceived as less desirable than either a human- or, in some cases, an AI-only decision maker.

KEYWORDS

Collaborative Human-AI Decision Making, Higher Education, User Perception

Introduction

In Higher Education (HE), Artificial Intelligence (AI) has the potential to enrich and extend learning outcomes and experiences, and to reduce the administrative workload for educators, by optimising teaching resources and providing algorithm-driven marking and assessment (Zawacki-Richter et al., 2019). However, the emergence of AI is an emotive and polarising topic, and its adoption may be perceived as de-valuing or displacing humans. Moreover, education presents a unique context in which trust can be fragile, and institutions must therefore ensure clarity and transparency in decision-making. In practice, *perceptions* of AI decision-making, regardless of the actual quality of the decision outcome, can significantly influence confidence and trust, and result in poor acceptance or rejection of the system (Sundar & Nass, 2001). AI has traditionally been associated with big data, statistics and machine learning, and this enables efficient, optimised, and data-driven decision-making. AI is also perceived as more rational and less emotional than people (Waytz and Norton, 2014). In a HE context, AI decision-making may thus be more suited to tasks that require 'mechanical' or algorithmic skills (e.g. timetabling and admissions, in which large datasets are managed and optimised). In contrast, tasks in which the decision maker is required to make subjective and intuitive judgements, or to understand and express emotions or navigate difficult social situations (e.g. marking written or creative work or resolving interpersonal conflicts), may be more suited to a human decision-maker (Lee, 2018). In practice, this distinction may be somewhat blurred, with many tasks requiring elements of both skillsets. It has therefore been suggested that a potential solution is to harness the complementary strengths of human and AI within collaborative decision making, in which AI does the 'heavy (data) lifting' and a human expert subsequently checks or refines the outcome (Dolgikh & Mulesa, 2021). Indeed, *performance* successes have been reported with such an approach, but it is unclear how this is *perceived* by students and educators, which could ultimately affect adoption.

In the first investigation, we explored attitudes towards human-only and AI-only decision-making in an HE context using both a survey (n=94, comprising students and academic staff) and two focus groups (each comprising 4 survey respondents) to determine whether attitudes differed based on the nature of the task (i.e. tasks requiring mechanical compared to human skills). Building on the findings from the first survey, we subsequently conducted a second survey (n=75), in which we maintained the same task distinctions and academic setting but offered a third option, notably, a

human-AI collaboration. In both surveys, which were hosted on https://www.prolific.com, participants were presented with four scenarios: admissions (academic and holistic assessment of candidate), marking (multiple-choice exam), course scheduling, marking (written essay). Inspired by Lee (2019), these differed in the degree to which they involved 'mechanical' and 'human' skills. In the first study, participants were told that for each scenario, the decision outcome was provided by either an experienced/expert human professional (e.g. administrator, professor, admissions officer, as appropriate) or an AI system, in isolation. In study two, a third option was added, in which the decision was made by AI and then "tweaked", "reviewed" or "adjusted" by the relevant human expert (Human-AI collaboration). Participants were asked to rate the fairness/accuracy, trustworthiness and emotional response/satisfaction associated with each scenario/decision maker, using 7-point Likert scales, where 1 indicated the least positive rating (typically, 'not at all'), and 7, the most positive (typically, 'completely', although the precise nomenclature and scale anchors depended on the factor under evaluation). Ratings were compared using one-way ANOVAs.

Results from study one show that AI decision-makers were perceived as fairer and more trustworthy than human decision-makers in tasks requiring ***mechanical skills***; emotional response was similarly high for both AI and humans. Higher ratings of fairness and trust were attributed to the efficiency and objectivity of the AI system when executing mechanical tasks, in which the decision-making process relied upon analysing facts and data against predefined rules and algorithms. In contrast, human decision-makers were considered to be susceptible to subjective emotions and personal preferences that could affect outcomes. Nevertheless, while participants stated that AI offered fairness and was "immune to exhibiting discrimination", it was also highlighted that human decision-makers offered "more profound insight" and could adapt their decision-making to take into account unexpected human factors that might only come to light when a decision is implemented. The positive emotional response associated with AI decision-making reflected the belief that AI exhibited "consistency, efficiency and impartiality" and its decisions were "transparent". Surprisingly, in tasks requiring ***human skills***, participants also perceived AI decision-makers as fairer than human decision-makers, and also equally trustworthy/reliable, although the emotional response towards AI was less positive for these tasks. This outcome reflected the potential for unfair or biased decision-making in situations where humans made the decision, for example, the "potential for nepotism when deciding on admissions" . However, it was also recognised that human decision-makers have a sense of responsibility and accountability, especially when problems or mistakes occur, and that the same accountability could not be attributed to AI as it lacked agency. Notably, this did not preclude the concern of algorithmic bias during AI decision-making, or indeed, any technical limitations (for example, in making 'value' judgments, see: Araujo et al., 2020), but it was felt that if any issues arose with an AI decision, it would be "transparent" and could be reviewed and corrected in a fair and impartial manner. As such, participants were generally supportive of algorithmic marking or qualitative AI assessment, for example. However, there was a notable concern expressed that students would feel "profoundly disrespected" if their written assignment or application was evaluated solely by AI and was never seen by a "fellow human". This supports the notion of a hybrid approach, in which an AI decision could be checked or reviewed by a human. However, results from study two suggest that ***collaborative Human-AI decision-making*** was generally considered to be less accurate, less trustworthy and likely to provide lower satisfaction in the decision outcome, compared to a human decision-maker, and in some cases, also to AI in isolation. This is perhaps surprising given the recommendations and expectations from study one, and, indeed, the performance benefits this hybrid approach purportedly offers (Dolgikh & Mulesa, 2021) and suggests that in an HE context, at least, a hybrid approach may fall short of its 'best-of-both' ambition. Results may reflect a lack of understanding of precisely how human and AI could work collaboratively in a HE context, but also reinforce the importance of determining how the decision-making process is perceived, not only how well it performs, particularly by those people who are directly impacted by the outcome.

References

Araujo, T., Helberger, N., Kruikemeier, S. and De Vreese, C.H., 2020. In AI we trust? Perceptions about automated decision-making by artificial intelligence. AI & society, 35(3), pp.611-623.

Dolgikh, S. and Mulesa, O., 2021, September. Collaborative Human-AI Decision-Making Systems. In IntSol Workshops (pp. 96-105).

Lee, M.K., 2018. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. Big Data & Society, 5(1), p.2053951718756684.

Sundar, S.S. and Nass, C., 2001. Conceptualizing sources in online news. Journal of communication, 51(1), pp.52-72.

Waytz, A. and Norton, M.I., 2014. Botsourcing and outsourcing: Robot, British, Chinese, and German workers are for thinking—not feeling—jobs. Emotion, 14(2), p.434.

Zawacki-Richter, O., Marín, V.I., Bond, M. and Gouverneur, F., 2019. Systematic review of research on artificial intelligence applications in higher education–where are the educators?. International Journal of Educational Technology in Higher Education, 16(1), pp.1-27.