Doctor Who?: Norms, Care, and Autonomy in the Attitudes of Medical Students Toward AI Pre- and Post-ChatGPT

Andrew Prahl1 and Kevin Tong Weng Jin2

1 Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore
2 Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

Abstract

This study adopts the combined TAM-TPB model to investigate attitudes and expectations of machines at a pre-career stage. We study how future doctors (medical students) expect to interact with future AI machinery, what AI usage norms will develop, and beliefs about human and machine autonomy. Semi-structured interviews were conducted. Wave one (N = 20) occurred 6 months prior to the public release of ChatGPT; wave two (N = 25) occurred in the 6 months following. Three themes emerged: AI is tomorrow, wishing for the AI ouvrier, and human contrasts. Two differences were noted pre-versus post-ChatGPT: (1) participants began to view machinery instead of themselves as the controller of knowledge and (2) participants expressed increased self-confidence if collaborating with a machine. Results and implications for human-machine communication theory are discussed.

Keywords: ChatGPT, health care, human-machine communication, generative artificial intelligence, GenAI

Acknowledgment: We would like to acknowledge the support of the Nanyang Technological University Undergraduate Research Experience on CAmpus (URECA) program.
Introduction

In the bright bluish light of the theatre, a group of humans form a circle in the middle. In deep concentration, one raises a blade and a voice pierces the silence: “Have you remembered to administer prophylactic antibiotics?” While the surgeon and their team in this operating room theatre are real, the voice is distinctly inhuman. “Yes,” the surgeon replies. This important reminder is not the last contribution to the operation that the surgeon’s robotic assistant will make. After an incision is made, the robot is perfectly positioned to retract the skin, giving the surgeon the best view of the viscera beneath. Along the way, there are constant updates about the patient’s blood pressure, volume of blood loss, heart rate, and oxygen saturation. Throughout, the assistant utilizes artificial intelligence (AI) to comprehend what is happening, adjust as needed, and provide constant reminders the human team needs to ensure that nothing is missed.

Though this scenario might seem futuristic, as if taken from the year 2100, a similar series of events may be seen in operating theatres across the world sooner than we think. And the operating room is just one place where machines are becoming more common in health care. Nearly all levels of care, from the family physician’s office to the most complex lab analyses are being constantly transformed by new technology. However, in contrast to the past’s relatively steady proliferation of medicine technology, the newest machine capabilities—heralded by some as the “4th Industrial Revolution”—are ushering in an age of exponential change (Evans, 2019). Recently, the widely known ChatGPT web application became the fastest growing app ever, outpacing the userbase growth of the former champions (and still tremendously popular) Instagram and TikTok (Hu, 2023; UBS, 2023). Clearly, when they graduate, medical students of today will be entering a world in which machines, especially those using AI, play an outsized role that will continue to grow throughout their careers.

It is imperative that we explore the attitudes of these future doctors toward their future machine companions. Ultimately, the older generations of medical professionals who began their careers without the aid of advanced machine tools such as decision-support systems or autonomous surgical robots will retire from the workforce. The new generation of medical professionals, much like the way younger generations grew up immersed in digital technology, are now emerging into their careers in an era dominated by advanced AI tools. But what will this relationship between doctors and machines be like? What clues can we gain now as to future sources of harmony between human and machine—or, less optimistically—can we identify future sources of tension? An extensive body of research in the human-computer interaction (Gibson et al., 2020), human-automation trust (Hoff & Bashir, 2015), and human-machine communication (Guzman, 2020) disciplines testifies to the importance of pre-existing attitudes and expectations when interacting with new technology. Though most medical students are not explicitly trained in AI, the ubiquity of articles regarding AI in popular media and AI’s overall salience in the public sphere suggests attitudes could be changing rapidly (Bartholomew & Mehta, 2023). We investigate these attitudes and beliefs here; we specifically focus on attitudes toward AI because it is most representative of cutting-edge and future technology.

In our investigation, we contribute to human-machine communication literature by stepping further back into the timeline of professional attitudes. That is, attitudes are formed
not only in the workplace during or after machine implementation (where most current research is situated), but also before the people enter the workforce. This study asks participants to be more forward-looking than the majority of extant research; in envisioning their careers, medical students must think in terms of decades. To guide our research, we adopt a theoretical framework that is suited to this extended temporal scope: the combined Technology Acceptance Model and Theory of Planned Behaviour (TAM-TPB) (Taylor & Todd, 1995), and we conduct two waves of our study, one preceding the release of ChatGPT and one after the release. Resultingly, our research provides insight into (1) the openness to using AI in their future careers, (2) the perceived usefulness of AI technologies, (3) the norms that medical students believe will form regarding the use of AI in the workplace, (4) the beliefs about personal autonomy in using AI, and (5) how these attitudes may have changed following the release of software that revolutionizes human-machine communication. We first review literature on worker attitudes to AI more broadly, previous work with medical students, and the TAM-TPB framework.

Literature Review

Artificial Intelligence (AI) is generally understood to refer to the ability of the robots or machines to perform higher order human cognitive functions, to “think” and “act” like humans (IBM, 2020). In recent years, AI programs such as Deep Blue and AlphaGo have come to the forefront of attention by defeating the brightest human minds in games such as Chess and Go. At the same time, some remain wary of machines’ ability to replace humans, adamant that “AI will not replace humans overnight” (Toews, 2021). For our research, we defer to the most widely used definition of AI as “agents that receive percepts from the environment and perform actions” (Russell & Norvig, 2009, p. 10), a definition that allows for the study of a diverse array of machines. Recognizing this broad application of AI, it is critical to understand its impact on diverse domains, especially in the context of human work. Given the issue of human replacement by AI is so salient, much previous work on perceptions of AI in the workplace has focused on this question. Our inquiry, however, takes a different angle to explore the understudied perspective of how future professionals visualize their potential co-existence with AI in the workplace. This shift in focus does not neglect the question of replacement but instead attempts to build upon it by addressing the envisaged human-machine communication in the professional landscape, specifically from the vantage point of professionals entering their field.

Worker Attitudes

Workers are not always welcoming to AI. Job insecurity and psychological distress are associated with fears about AI taking over jobs for workers in a Philippines call center (Presbitero & Teng-Calleja, 2022). Similarly, concerns arise among North American pharmacists, suggesting apprehension across many job types. These professionals may experience “automation anxieties” (i.e., concerns about job loss) when confronted with the advance of machines into their professions. AI may threaten the identities of employees, especially as it relates to job functions, and may pose a risk to the “social fabric of work” (Selenko et al., 2022, p. 1). This is generally negative not only for employees, but for organizations as well because
job satisfaction (or lack thereof) affects trust between employee and employer (Rich-
ter & Näs­wall, 2019). Media coverage of AI is widespread, and pessimistic views of AI
are commonplace (Siegel, 2019; Sun et al., 2020). Even if presented alongside optimistic
views regarding AI and labor, the tendency of humans to give more weight to negative
outcomes is well-documented (Kahneman & Tversky, 1979). From this perspective, our
study acquires an important dimension: it attempts to gauge the pre-existing sentiments of
people just venturing into the workforce. We are interested in assessing whether the preva-
luent AI narrative, often tinged with negativity and uncertainty, has influenced their outlook
on working with machines. The public release of ChatGPT in November 2022, given its
rapid adoption—setting the record for fastest application ever to reach 1 million users (Hu,
2023; UBS, 2023)—along with its accessibility and widespread media coverage, is an ideal
moment to study. Hence, our research design is crafted to capture the sentiment backdrop
that the participants might bring to the table.

Leveraging the vantage point of the health care industry benefits our study. Health care
is one field in which the application of artificial intelligence is rapidly growing. A number
of studies have found varying support for AI in medicine; however, the perception that
patient privacy must be protected and that clinicians should be the main participant in
decision-making with patients is nearly universal (Scott et al., 2021). Examples cited include
the ability of AI to interpret diagnostic imaging, aiding in pharmaceutical development
and streamline administrative tasks (Shah & Chircu, 2018). This prioritization of human
involvement is not ubiquitous across all industries, with aviation being one notable exam-
ple where machine autonomy is more embraced. In essence, this distinction underlines
that despite technological advancements, the human element is considered central to the
practice of health care.

Medical Students and Machines
In recent years, studies have been performed to find out more about medical students’ per-
spectives on Artificial Intelligence (AI) in health care, often with a particular emphasis on
the choice of residency or speciality, as well as how it might change the role of doctors in
the future. This has been done in countries such as the UK (Sit et al., 2020), the US (Park et
al., 2021) and Canada (Gong et al., 2019). Across the board, students surveyed felt that AI
will play an important role in health care in the future, and in one study, 89% of students
stated that teaching in AI would be beneficial for their future (Sit et al., 2020). Cho et al.
(2021) found that although there was interest in AI among medical students, there was a
discrepancy between degree of interest and concrete AI education. The primary motivation
for many studies of medical students is for education/curriculum planning purposes, and
is driven by questions related to how AI should be integrated into medical education. It
may not be surprising to see that most of the medical schools which the survey participants
came from did not have AI-related content in their syllabi. However, it may be surprising to
learn that the vast majority of medical education programs do not include AI in any form.
More broadly, other than simple “how to use X machine” education that students receive
in experiential settings, medical education is devoid of education about human-machine
communication, despite the machines in the workplace being widespread.
The curricular-focus of research on medical students has also aimed to address what specialties students believe will be most affected by AI. Two fields purported to be most susceptible to being replaced by AI were radiology and dermatology (Cho et al., 2021; Gong et al., 2019). Many students felt that radiology would be the specialty most affected by AI, with around half of the students surveyed stating that their interest in radiology had been negatively impacted by the development of AI. But the more daunting question of if AI has made the medical profession less attractive in general remains unanswered. Our work will provide insight as attitudes toward AI in professional life are inevitably a combination of both how people feel about AI and how they feel about the profession itself. Both research with medical students and in medicine generally have provided little theoretical basis for understanding attitudes toward AI. This is unfortunate as medical students in particular are in an ideal position to inform theory about how workers approach AI. Medical students are at the very cusp of long careers in medicine. An investment in medical school is a serious undertaking that prepares one for a particular industry; if AI takes over a doctor’s job, they cannot just jump to a different industry and apply their skills in logistics or accountancy. Like many other areas of skilled labor, the “I’ll just do something else” approach may not be viable given the limited transferability of skills. Furthermore, while they are forced to think deeply about their careers, medical students are not yet embedded in real work environments where their attitudes toward AI may be jaded by bad experiences or the dysfunctional work environments that have plagued previous attempts to introduce technology in organizations (Stam et al., 2004). This study provides a novel investigation into the attitudes of medical students as we investigate how the release of a disruptive technology (ChatGPT) may cause shifts in how medical students think about machines.

**Perceptions and Expectations**

In investigating perspectives that are relatively unadulterated by work experience, our work follows previous investigations in the human-machine communication tradition (e.g., Guzman, 2020; Simmler et al., 2022). Our sample combines the best advantages of both professional and general population samples. For medical students, AI carries strong professional implications but they are not yet working with AI every day. To maximize the advantages of our sample our investigation is guided by the combined Technology Acceptance Model and Theory of Planned Behaviour (TAM-TPB) (Taylor & Todd, 1995). We chose TAM-TPB because it reflected the qualities of our sample, because it includes items geared toward general populations/issues from the widely used Theory of Planned Behavior (Ajzen, 1991) and because it also includes professional focused items from the similarly popular Technology Acceptance Model (Davis, 1989).

TAM-TPB proposes that attitudes toward technologies are determined from practical factors one would encounter in the workplace such as the perceived usefulness of the technology or the ease of using it. But attitudes also originate from the norms people perceive regarding technology. For instance, medical students’ perception of AI in prognosis management would influence their attitudes. Further, as their careers progress, their attitudes toward AI will be shaped by both professional, patient, and societal expectations. Norms are important both on the professional and public sides of health care: providers may have
one set of expectations about how machines should be used, but the public (patients) may have different expectations; these differing expectations must be negotiated. TAM-TPB also proposes that one’s *behavioural control* can affect intentions to behave. *Control* has multiple meanings in the workplace. While the straightforward interpretation is related to self-efficacy, control may also manifest in company policies, directives from superiors, or demands from patients that a technology be used or prohibited. Or, machines may demonstrate better performance to humans who resulting feel inferior and compelled to defer to machines. Behavioral control therefore is related to the struggle for autonomy that is frequently discussed in studies of human and machine (Fortunati & Edwards, 2021; Schaefer et al., 2016). In light of these considerations, adopting the TAM-TPB framework enhances our study as it encompasses both individual and external factors that contribute to one’s intention toward new technologies. In sum, the use of the TAM-TPB leverages the strengths of our sample and provides structure for our contribution to human-machine communication theory. As such, we are guided by three research questions:

**RQ1:** How do medical students foresee the usefulness of artificial intelligence in their careers?

**RQ2:** What future norms regarding the use of artificial intelligence in medicine do medical students foresee?

**RQ3:** How do medical students envision personal and professional autonomy in relation to the use of AI in their future workplaces?

### Method

A total of 45 (N = 45) medical students participated in semi-structured interviews. The first wave of interviews (n = 20) were collected between December 2021 and June 2022. First wave participants were sourced initially through recruitment posters (n = 8) and snowball sampling (n = 12) by asking participants to recommend others for the study. In the second wave (n = 25), collected between January 2023 and June 2023, fewer students (n = 5) responded to recruitment posters with the remainder sourced through snowball sampling (n = 23). Recruitment in the second wave differed from the first in that there were two research team members independently snowball sampling in order to mitigate the selection bias risks inherent in snowball sampling. The interviewees selected were medical students in their second to fourth year of medical school from two undergraduate medical schools in Singapore. Ages ranged from 19 to 24. Notably, Singapore’s medicine education program is compressed compared to the United States where equivalent schooling is often pursued after obtaining a bachelor’s degree. Thus, our sample is 3–6 years younger than similar studies conducted in like education programs elsewhere.

Interviews were conducted using a mix of face-to-face and video call mediums because some students were under isolation directives resulting from COVID-19 mitigation measures. These directives were loosened between wave one and wave two, meaning more interviews were conducted in-person in wave two (56.5% in-person) than wave one (40.0% in-person). Interviews were conducted by study team members who were fellow medical students in Singapore. Research assistants were earning course credit for participating in
an undergraduate research program; participants were not compensated. The interviews adopted a semi-structured format. A pre-defined list of questions was prepared, but conversations were allowed to flow naturally based on interviewees’ responses. The only difference between the wave one and two interview guides is that wave two included allowances for participants to go on tangents related to the recently released (with much fanfare), generative AI tools (e.g., ChatGPT), but we did not change any questions in the interview guide to ask about generative AI specifically. Questions were structured around the combined TAM-TPB model, although posed in layperson language to encourage more natural discussion. Some questions were general conversation openers (e.g., “Do you see AI changing the role of doctors during your career?”) whereas others were more targeted at concepts of interest such as fear (e.g., “How do you think you would feel if your boss came to you and told you that you must use a new AI tool because studies show it has better judgment than you?”). Interviews lasted between 25–45 minutes.

Coding and Content Analysis

Interviews were transcribed, analyzed and coded. Our approach follows the reflexive thematic analysis method (Braun & Clarke, 2006; Byrne, 2022) with the coding framework being developed and evolving as more interviews were being transcribed and analyzed. A framework of core themes was built based on the analysis of the initial transcripts using an inductive approach. This framework was built upon and modified as more and more transcripts were coded and analyzed. Earlier interviews were then revisited and re-coded in an iterative process (Braun & Clarke, 2006). In wave one, after completion of the initial coding scheme, study team members met to begin mapping the codes to concepts in the TAM-TPB model when applicable. This process, which we favored over a more structured deductive approach, can result in the identification of codes that fall outside the conceptual framework; we were agreeable to this due to exploratory, futuristic nature of our research questions. But these codes may be problematic from a theoretical view if they are over-fit into the framework. To avoid this conundrum, a crucial step was the consensus-building discussion among our study team. After discussion, the study team agreed that all codes that could not be mapped to a construct(s) in the TAM-TPB were related to student’s educational curriculum (which was an inevitable topic of conversation given the interviewer was a peer student). Thus, we place these codes in an education category. All codes and frequencies are shown in Table 1. In wave two, the process differed in that the study team already had the codes developed by the study team members in wave one. The research assistants in wave two were supplied with the codebook, and coded transcripts from wave one. Wave two research team members then coded five transcripts from wave one, and one research team member from wave one re-coded the same transcripts. Intracoder reliability for the wave one and wave two team members was 96% agreement. For reliability between all three wave two members, we elected to calculate Krippendorf’s alpha to mitigate problems with percentage agreement (Hayes & Krippendorff, 2007); reliability was acceptable $\alpha = 0.882$. The only source of significant disagreement in wave two data was how to cover discussions of generative AI and ChatGPT, which was not accounted for (the technology did not exist yet) in wave one. We place these codes into a “new technology” theme which is listed separate from the themes present in wave one and two (see Table 1).
TABLE 1  Codes, Frequencies, and Theme Mapping

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Percentage of All Codes</th>
<th>Percent of Interviews Code Occurred</th>
<th>Theme</th>
<th>Sub-Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration</td>
<td>50</td>
<td>8.25%</td>
<td>100.00%</td>
<td>AI is Tomorrow</td>
<td>AI will be Demanded</td>
</tr>
<tr>
<td>Patient Desires</td>
<td>85</td>
<td>14.03%</td>
<td>100.00%</td>
<td>AI is Tomorrow</td>
<td>AI will be Demanded</td>
</tr>
<tr>
<td>Growth of Technology</td>
<td>68</td>
<td>11.22%</td>
<td>100.00%</td>
<td>AI is Tomorrow</td>
<td>AI-Saturated Work</td>
</tr>
<tr>
<td>Ubiquity</td>
<td>15</td>
<td>2.48%</td>
<td>33.33%</td>
<td>AI is Tomorrow</td>
<td>AI-Saturated Work</td>
</tr>
<tr>
<td>Inevitability</td>
<td>60</td>
<td>9.90%</td>
<td>100.00%</td>
<td>AI is Tomorrow</td>
<td>AI-Saturated Work</td>
</tr>
<tr>
<td>Optimism</td>
<td>34</td>
<td>5.61%</td>
<td>75.56%</td>
<td>AI is Tomorrow</td>
<td>Tomorrow is Better</td>
</tr>
<tr>
<td>Safety</td>
<td>20</td>
<td>3.30%</td>
<td>44.44%</td>
<td>AI is Tomorrow</td>
<td>Tomorrow is Better</td>
</tr>
<tr>
<td>Pessimism</td>
<td>18</td>
<td>2.97%</td>
<td>40.00%</td>
<td>AI is Tomorrow</td>
<td>Tomorrow is Better</td>
</tr>
<tr>
<td>Humans: Skill Comparison</td>
<td>12</td>
<td>1.98%</td>
<td>26.67%</td>
<td>The Human Contrast</td>
<td>Cognition and Intuition</td>
</tr>
<tr>
<td>Humans: Positive Comparison</td>
<td>15</td>
<td>2.48%</td>
<td>33.33%</td>
<td>The Human Contrast</td>
<td>Cognition and Intuition</td>
</tr>
<tr>
<td>Useful as Assistant</td>
<td>12</td>
<td>1.98%</td>
<td>26.67%</td>
<td>The Human Contrast</td>
<td>Different Colleagues</td>
</tr>
<tr>
<td>Coworkers</td>
<td>10</td>
<td>1.65%</td>
<td>22.22%</td>
<td>The Human Contrast</td>
<td>Different Colleagues</td>
</tr>
<tr>
<td>Ease of Profession</td>
<td>14</td>
<td>2.31%</td>
<td>31.11%</td>
<td>The AI Ouvrier</td>
<td>A Better Professional</td>
</tr>
<tr>
<td>Choice of Use</td>
<td>12</td>
<td>1.98%</td>
<td>26.67%</td>
<td>The AI Ouvrier</td>
<td>A Better Professional</td>
</tr>
<tr>
<td>Health Care Industry</td>
<td>18</td>
<td>2.97%</td>
<td>40.00%</td>
<td>The AI Ouvrier</td>
<td>A Better Professional</td>
</tr>
<tr>
<td>Ease of Use</td>
<td>7</td>
<td>1.16%</td>
<td>15.56%</td>
<td>The AI Ouvrier</td>
<td>Taking Difficult Work</td>
</tr>
<tr>
<td>Ease of Workload</td>
<td>18</td>
<td>2.97%</td>
<td>40.00%</td>
<td>The AI Ouvrier</td>
<td>Taking Difficult Work</td>
</tr>
<tr>
<td>Useful as Tool</td>
<td>40</td>
<td>6.60%</td>
<td>88.89%</td>
<td>The AI Ouvrier</td>
<td>The AI Toolbox</td>
</tr>
<tr>
<td>Humans: Negative Comparison</td>
<td>39</td>
<td>6.44%</td>
<td>86.67%</td>
<td>The AI Ouvrier</td>
<td>The AI Toolbox</td>
</tr>
<tr>
<td>Education</td>
<td>15</td>
<td>2.48%</td>
<td>33.33%</td>
<td>Education</td>
<td>Education</td>
</tr>
<tr>
<td>Electives</td>
<td>9</td>
<td>1.49%</td>
<td>20.00%</td>
<td>Education</td>
<td>Education</td>
</tr>
<tr>
<td>Generative AI (e.g., ChatGPT)*</td>
<td>35</td>
<td>10.78%</td>
<td>100.00%</td>
<td>New Technology</td>
<td>New Technology</td>
</tr>
<tr>
<td>Total</td>
<td>606</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Code only occurred in wave two, percentages calculated for wave 2 data only.
Thematic Development

After the study team reached sufficient agreement on codes, the process of thematic analysis began with building upon the initial content coding. First, codes were categorized and abstracted into larger conceptual themes. Through discussion, themes and potential subthemes were identified. Ultimately the study team settled on three themes with a total of ten sub-themes that accurately summarized the data. Given the format of the interview questions that were structured around the combined TAM-TPB model, themes unsurprisingly gathered around several constructs in the model. We then wrote the results section in a collaborative manner to ensure agreement on content.

Results and Discussion

Corresponding with the interview guide which used the combined TAM-TPB framework, and the research questions regarding concepts in the model, our three main themes largely corresponded with the perceived usefulness, perceived norms, and behavioral control. However, in constructing our themes we drew relevant information from across the entire conversation instead of only responses to specific questions. Thus, we discovered new angles on what these oft-used concepts mean given the unique perspective of our participants. We first present the three main themes that occurred across both waves. Then, we discuss differences between wave one and wave two.

AI is Tomorrow

The inevitability of AI in medical workplaces was a constant theme in interviews, so much that some codes and sub-themes (e.g., AI being demanded by administration and patients in the future workplace; AI use continuing to grow) occurred in 100% of interviews. Sentiment toward this future demand was not universal, with participants noting both pros (e.g., improved safety) and cons (e.g., need to retrain frequently). Regardless, no respondents saw a future without AI in the workplace, nor did any suggest that there will be large amounts of resistance to this change. Despite acknowledging the changes that AI may bring, respondents largely believed AI wouldn’t change what it means to be a doctor. Some students felt that AI would not change the role of doctors very much, at least not in the near future. One student stated, “I think the fundamental roles of the doctor will still be there and won’t be completely replaced” [14]. Another said, “I don’t think AI will drive us out of business . . . down [that] path, hopefully AI is more friend than foe” [12]. There is a belief that AI is nearly synonymous with the workplaces the respondents will work in throughout their careers: AI is not a “maybe” in the workplaces of tomorrow; AI is tomorrow. And these workplaces will become further saturated with AI over time. Respondents understood that AI capabilities will increase, “you know as AI gets more precise and more experience, higher datasets, it can even go up the ladder and take up more specialised skills” [4]. While these possibilities are acknowledged, there is still uncertainty about just what AI will be capable of, and how quickly it will move, as summed up by one respondent: “AI will move faster than we expect but also slower than we expect, if you know what I mean” [35].
Addressing our first research question, we find a belief that the workplaces of tomorrow may be improved by AI, “I will be welcome to this sort of changes” said one respondent, continuing, “If it helps to make your job more efficient and more error-free, then I think probably it would be nice to welcome that sort of change” [7]. But, a less optimistic view was present in roughly 40% of interviews. For example, AI may bring problems to the future workplace, especially in the implementation side that is plagued by long timelines and testing, “They will experiment, they will trial, they will roll out a pilot programme, and that’s all fine and good, but for mass adoption, that will take a long time” [41]. In addition to dealing with long rollouts, respondents are not enthusiastic about the potential ethical questions introduced by AI:

“Then there’s also the whole legal aspect, like what if a patient is misdiagnosed? Computer says it’s a correct diagnosis, like the answer key says it’s correct, but in reality it was misdiagnosed like based on autopsy or tissue biopsy or too late, then it’s a dispute, like why did the doctor go against the decision of the computer, how valid is the clinician’s experience and what not.” [13]

Nevertheless, respondents see patients as being open to AI, perhaps even demanding its use from health care systems. Safety improvements may prompt this, “the overall impact [of AI] will be positive for the patient, then I think that I will also be glad to accept that change because I mean, I will accept the fact that probably we do make judgement errors” [2]. But this demand may follow generational shifts, “Probably the older generation [patients] will still prefer the more personal, they probably rely less on technology compared to the younger generation who might have differing thoughts on AI” [33]. Some respondents also offered a countertheme to this, suggesting patients will ultimately be unhappy, “. . . if you put AI’s formulaic way of thinking into the practice of medicine, then you would have a lot of unhappy patients and a lot of hurt patients” said one respondent [11].

All of these thoughts come with the assumption that AI will be adopted and will saturate the workplace. In turn, respondents shift their thinking from the traditional “perceived usefulness” as being a measure of likelihood of use. Rather, usefulness is a maybe but the presence of AI is not. And in truth, respondents saw usefulness as only a weak maybe. Bar- ring a few pessimistic thoughts, the majority of comments saw this AI tomorrow as a better place, especially in reference to newer generative AI technologies (e.g., ChatGPT) present in the second wave of data collection. This answer to our first research question was reflected in our data, too, with optimistic futures codes occurring more than pessimistic ones. The results also provide some insight into our second research questions about norms. AI is the norm tomorrow; this AI tomorrow is inevitable and good.

**Wishing for the AI Ouvrier**

If AI is an inevitable fixture on doctor’s careers, they want AI as a sidekick instead of a substitute. It is telling that the most frequently occurring code in this theme—the idea that AI is only a tool—occurred in nearly every interview (88.89%). The right tools will make the workplace more pleasant, as indicated by the discussion of AI’s potential to reduce workloads (40%) or ease other aspects of the profession (31.11%) provided that users will be
able to choose when and how to use AI (26.67%). One respondent recounted the benefits of AI they have heard of in other industries, “Not so much in the medical field, but in other applications like industrial applications, how it’s used to streamline predicting your supply chain demands, for business analytics” [5]. AI poses a threat when it acts as a substitute,

“I guess if [the AI] something that like doesn't take away the joy of practising in that specialty, then I wouldn't mind. The converse example is if I want to do surgery, but the AI is doing everything all the surgeries, then I don't think I would want to do it. It takes away the joy of actually using your own hands and operating on the patient.” [6]

This quote deserves closer inspection: joy is under threat because it removes some degree of interaction. But the imagery of “actually using your own hands” is meaningful as well—people do things with our hands, it is the literal feel of the job. The removal of this physical connection with the job may represent a disembodying prospect for doctors. But this fear does not manifest itself in a resistance to AI or a decrease in perceived usefulness. Instead, people wish to carefully pick specialities to avoid this fate. Again, we see the implication that AI is inevitable. One cannot resist an unstoppable force.

Respondents discussed the norms (RQ2) they foresee around AI use. The norms were often hopeful, respondents wish that AI will “serve as an adjunct,” [7] “provide second opinions,” [36] or, as one respondent candidly admitted their fallibility, AI can be a backup, “We do make mistakes, the AI could back us up” [40]. One respondent represented the thought of AI as a laborer for those tasks that are either time-consuming or difficult. The respondent pointed out the task of looking at diagnostic images, “I think what we're all looking forward to is skipping the whole interpreting patho slides part, it would be amazing!” [9]. Another respondent expressed a similar thought bluntly: “If you’re talking about things that are more like clinical and diagnostic, like pattern and image recognition, then ya, even better AI does it, because we all suck at x rays and that kind of thing” [12]. AI not only makes up for the shortcomings of humans but takes up the tasks that doctors do not enjoy. Ouvrier, an obscure English word adopted from French, meaning “a workman who is employed to do heavy work requiring little skill” (Cambridge, n.d.), describes these hopeful norms well. AI can take difficult work, it can make participants better professionals, but AI sits below humans on the work-value chain—a sentiment that is perhaps exaggerated due to the norms and social prestige that come with being a doctor. While doctors still make leadership decisions and do hands-on work, it is expected of doctors to happily pass off undesirable to their ouvrier colleagues: AI.

The Human Contrast

Our third research question probed the issue of autonomy as a potential source of tension between human and machine. Through the lens of TAM-TPB, this is a question about the control that humans will have over AI: What things will humans always be best at, what things will AI instead be in the driver’s seat? In our interviews, this discussion elicited many comparisons of human and machine qualities, typically involving comparisons of machine cognitive skills compared to human intuition (26.67%). We were surprised that the newest
generative AI technologies present in wave two did not change the nature of the human versus machine comparisons participants discussed. For example, participants believe that AI cannot control relationships with patients, including via the physical touch hinted at by other respondents above, “the personal connection with your patient, stuff like physical exam and things like that, I think those are things that AI cannot replace” [6]. Being a human health care provider is not only about touch, humans also provide assurance. One respondent put them in the position of patient, “As patients we go to a clinic to seek reassurance or to talk to someone about something you couldn't tell anybody, so probably I’ll still find a human doctor” [34]. Another echoed the sentiment:

“If it's something that's worrying me, something that is out of the ordinary, even if it's something that is quite straightforward for a doctor right, then I think I would want some reassurance, like if it's a machine, I don't think I would get a lot of reassurance.” [15]

One respondent offered a blanket statement about why humans still will remain the defining part of health care, “… in the end, healthcare is a service, and you still need the human touch as they call it” [8]. The unsaid implication similar to the issues present in other human-machine communication work, such as the need for machines to recognize emotion to be effective in health care (Kim et al., 2021). Or, in other words, can machines really provide the care part of health care.

Other respondents dismissed the idea that AI becomes controlling of human doctors because of the onerous requirements of creating effective AI systems, “I think if you want to implement an AI to that, it needs a s*** ton of data,” said one [3]. Another discussed human intuition,

“I think a lot of the work that doctors do is also very intuitive, for example they see certain signs and they're able to synthesise what’s going on based on their clinical acumen. I think it's probably going to take quite a while or even ever for a system to be able to synthesise that amount of knowledge, not just hard knowledge that you can feed into a computer but probably also some other kinds of soft cues.” [7]

Here, “hard cues” are the domain of the machine. But humans provide a softer touch; one that is out of touch for machines. But machines can also be superior to humans (33.33%). Many acknowledged the strengths that AI has which give it an advantage over humans. One respondent said, “If you have something that can process a million times for information than we can, and really start to see the patterns even we forget, [it can] make things more streamlined” [22]. Another alluded to the amount of health care research and evidence-based treatments possible, “I think our human minds really cannot fight that amount of information, and if it does get that far then that would be an invaluable tool” [24]. It is clear that the respondents understood that computers and machines have much greater “processing power” than human beings can muster, even if this does not manifest as control.
There were some limited places students could see AI put in the controller’s seat, mainly on the administrative side of health care. For example, “You’ll probably see it adopted earlier in the operational side of things, because that is the part where AI is already engaged, it’s all numbers, optimising bed slots, allocation of resources, so that’s probably where it’s gonna start” was the opinion of one student [5]. Another student talked about “Crowd control, or scheduling doctors for clinic appointments, like how many doctors you might need at a particular time of the year, if it coincides with say flu season and travel incoming, even operational things like that” [12]. While these limits on human control may primarily originate from management, “Whether they want to let AI have that much power . . .,” [21] respondents still foresee having control over the use of AI generally. It would be an exaggeration to say these predictions are made with full confidence, though. One respondent hinted at some nervousness, “[If] the research does show that AI is significantly better than a human at certain tasks then well, so be it. I’m not sure that day would be anytime in the near future, but that could be famous last words” [7].

Pre- and Post-ChatGPT Differences

Repositioning on the Timeline of Machinery

Our first theme, “AI is Tomorrow,” is in reference to many participants discussing AI as being the future, whereas them themselves are currently not in the future but are ready to learn. Thus, we could rename the theme, “AI is Tomorrow, I am Today.” This sentiment is best reflected in the wave one data. If working solely off the wave two data we may be tempted to rename this theme, “AI is Today, I am Yesterday.” It is difficult to express the distinct sense of discomfort that was palpable in some wave two interviews, especially when discussing the latest technologies. Every participant who mentioned ChatGPT discussed it as a user of the technology rather than just an observer. Several participants remarked how “good,” “fast,” “expansive,” and “revolutionary” the technology is. As is well-documented in recent surveys, participants were using ChatGPT as everything from a writer of class assignments, to a study-buddy and even (as admitted by two) a second-opinion or quick reference tool in the clinic (The Learning Network, 2023). The key difference here is not that there is an overwhelming sense of pessimism in wave two interviews, but rather a sense of resignation. Wave two participants seem keenly aware that AI systems of the future (and perhaps current) will give patients faster, more detailed and—importantly—better answers than they themselves can. One participant made a telling statement, “My Aunt is always asking me questions about her conditions like I know something, I am only student lor . . . but let’s say right now she ask me question, do I go to my own knowledge, look in a [text]book or something, or just use ChatGPT and see if it is a reasonable answer [pause], ChatGPT of course!” [43]. Another student spoke about ChatGPT’s communication skill,

“It is amazing. If I ask it questions about something like pain management I will prompt it to speak empathetically. I say speak like Brené Brown, it delivers a better answer than me. It is soft and caring, not like me.” [42]
Another said it quite directly, concurring with other participants who spoke of the sudden advance of AI technology, when they said that chatting with a computer just “went from zero to hero” overnight [23].

While these quotes reflect the sentiment well, it was also reflected in the participants’ evaluation of their education. While we elected not to have any themes specifically about education, when the topic arose in wave two the sentiment was distinctly more pessimistic. In wave one, participants wanted to learn more about AI but understood why they did not; by wave two the participants were frustrated, perhaps some even bitter, that they did not. Many participants spoke of the incredible rise of ChatGPT, coming out of seemingly nowhere and being “orders of magnitude” better than any previous technology for seeking medical information [39]. We believe this represents an overall sense of feeling behind technology. Before, participants are aware the technology is advanced, they are aware it is improving, worryingly so. But they still hold the keys to their success, they feel as if investment now in education will allow them to keep up with technology. In wave two, the situation feels more hopeless. AI has not just inched ahead, it has leapt ahead, and the gap is only growing. In turn, there is a clear sense of participants feeling that they are a past landmark on the timeline of AI’s advance, rather than the status quo of AI being a tool on the timeline of human advance.

**Better Together; Worse Alone**

We were especially intrigued by a number of participants who expressed what can best be described as lowered self-confidence after becoming familiar with ChatGPT. “It made me wonder ‘what am I doing here?’” said one participant after becoming familiar with the technology [34]. Another said, “I have no doubt this morphs into something that makes for better care, we all want that, but I also want to have something to contribute” [38]. To be transparent, we only discovered this theme about one third through the second wave interviews. This may have caused the interviewers to frame some questions or prompts differently, so we interpret this theme with caution. However, there is a silver lining here which is much more clear in the data than the sense of decreased self-confidence: increased machine and self-confidence. In other words, participants feel enlightened and enabled by ChatGPT, it unlocks opportunities they did not have before due to human constraints.

Perhaps this new sense of confidence in human-machine hybridity is best summed up by a participant describing her enjoyment of using ChatGPT, “Honestly I find it . . . invigorating! I really enjoy trying to find that answer and wording I am looking for; I may know what it knows but it knows how I want to say it” [35]. Notably, the context of this comment was the participant describing homework assignments where she is instructed to describe what you would say to a patient in given scenarios. Another discussed wishing to find work environments that facilitate the use of the technology, “It is the way, [telemedicine] always seemed nice and all you know less prep less buffer, but if I can use this it is a game-changer, I won’t be able to type fast enough in-person” [45]. We noted many similar instances of participants—enabled by AI—feeling more knowledgeable, more well-resourced (e.g., time), and far more confident in their communication skills. ChatGPT has ushered in a new self-confidence through human-machine hybridity.
Discussion

The rapidly evolving human-machine communication landscape continues to redefine our sense of human-machine communication—our study contributes to this understanding both in terms of theory and practice. Our research shows the utility of the combined TAM-TPB model for investigating the attitudes, perceived norms, and behavioral control which are antecedents to communicating with machines. Attitudes and expectations are a major theme in human-machine communication research (e.g., Dearing, 2021; Gambino et al., 2020), but studying the underlying componentry of them has been challenging. We adopt a qualitative approach similar to previous work investigating the underlying assumptions people have about machines (Guzman, 2020), but structure it with a model providing new perspective. We make several resulting contributions to human-machine communication theory.

Utility of Combined Theory

The study of human-machine communication can be approached from a multitude of theoretical and methodological backgrounds (Fortunati & Edwards, 2020). As such, we see that the use of a theoretical framework that combines two models is useful for discovering new perspectives, and the unique dynamics of human-machine communication can bring new meaning to oft studied concepts. The TAM-TPB model provided us with practical concepts such as usefulness and efficacy but also catered to the forward-looking nature of our sample, who are not professionals yet, by asking them to consider future norms. Our results provided a new angle on the notion of perceived usefulness, for example, is most frequently conceptualized as a cause for use or non-use of machines. Our research suggests the meaning of the concept shifts in situations where use of automation feels inevitable: usefulness becomes a proxy for the enjoyment of interacting with machines and their effects—positive or negative—on the work environment. This is an important theoretical consideration in future research given the pace that machinery is being adopted in the workplace and given that communicating with machines in the workplace is becoming less of “just an option” and more often a “compulsory” part of the workflow, regardless if the machines are thought to be useful (Bulchand-Gidumal, 2022, p. 18).

We also show how studying norms can encourage creative thinking and reveal implicit beliefs about machines. We find that asking people to envision future norms can elicit the hopes and fears of participants. When looking far enough forward, norms are infused with as much hope as they are with fact, and provide a window into uncertainties that people have regarding their future relationships with machines. Structuring our question around future norms provided a different approach to study uncertainty than that used in extant work that investigates workers already on the job (e.g., Piercy & Gist-Mackey, 2021). It is significant that our findings echo this work. Piercy and Gist-Mackey (2021, p. 191) found that pharmacy professionals can experience “automation anxieties,” for example. Our work suggests that these anxieties are not entirely a result of machines coming onto the workplace, but these anxieties may be present far before professionals enter the workforce at all.

People may become more confident in their abilities when paired with AI. Our findings in this area (salient post-ChatGPT) show that the conflict between human and
machine autonomy can be studied from a perspective of control. Our line of questioning derived from the concept of self-efficacy was especially well-tuned to exploring this phenomenon. Given that industry is perhaps only in the very beginning of a “Cambrian explosion” (Matsuoka, 2018) of big data and AI development, this dynamic will take on greater relevance as future disruptive technologies are introduced and redefine human-machine communication. Thus, the seesaw-like sense of individual confidence decreasing but individual-plus-machine confidence increasing—witnessed in our study—is a promising area for future research. The instances we recorded of participants mentioning communication specifically (e.g., discussing pain management), are especially intriguing, and the question is broader than just medical students and medical settings. Further research should be conducted on how confidence in communication skills (among other skills) is affected by the introduction, use, and expertise with new machine communication technologies.

**Temporal Dynamics of Machine Agency**

The dynamics of human and machine agency manifest in a unique way in our study. In describing our results, we leaned heavily on the notion of time and control. However—in interpreting our findings—it is a mistake to simply reduce time to a linear concept, with a before, during, and after. For example, just describing humans or machines as ahead or behind is not an accurate characterization because behind implies inferiority, which is not always reflected in our data. Rather, the emergence of ChatGPT has affected participants in that they are no longer in control of their own timelines in regard to technology. In wave one, the self is the reference point on the timeline of experience, skill, and knowledge. Post-ChatGPT, the machines are the reference point. Machines control the timeline and advance at will, humans remain stationary, bound by unchangeable cognitive, emotional, and time limits. Hence, humans surrender their timelines and now live on the timeline of machinery.

What emerges from this, perhaps, is a new structure of how professional knowledge, practice, and standards are set. Throughout history, there is no obvious challenge to humans being the standard-bearers in all of these domains. Hence, the accumulation of knowledge, for example, in medicine is determined by what people determine to be correct. But the emergence of ChatGPT, as just the first in an inevitable line of improving AI technologies, appears to be tearing down this human-controlled structure. It is reminiscent of the work of Gibbs et al. (2021) who contend that technological systems can create new structures in workplaces. However, what we witness here extends beyond the workplace and into the personal mentalities of future professionals. AI is not the relatively simple algorithms that are “continually produced and reproduced by human action,” or that “evolve in a recursive relationship with human actors” (Gibbs et al., 2021, p. 165). Rather, AI systems such as ChatGPT are on the cusp of a transformative era where they go beyond enhancing humans and become autonomous entities that redefine the pace of knowledge acquisition and application. In this future, AI pioneers the benchmarks of efficacy and efficiency in various professional domains, forcing humans to adapt to machines’ pace rather than shape machines to human needs. Therefore, our data suggests a paradigm shift in human-machine interaction. Historically, machines operated within the confines of their programming. The generative AI revolution is changing this. It is vital to grasp the nuance here: This isn't about
a machine's dominance over humans, nor is it about machines making humans obsolete. Instead, it's about machines setting a pace that humans struggle to match. Machines, largely a passive tool in the existing workplace, are transitioning dynamic communicators that redefine the contours of human expertise.

This shift has practical and theoretical implications. If the bar of professional excellence is set by a machine, then humans must chase ever-changing—perhaps even elusive—standards. It's not just about keeping up anymore; it's about continuously recalibrating one's knowledge and skills to synergize with machine capabilities. For human-machine communication theory, we see an interesting convergence point with Banks and de Graaf's (2020) notion of agency that is a foundational concept of their agent-agnostic model of transmission. Banks et al. describe agency as the capacity to make a difference through action (p. 28). Multiple participants in our study mentioned that ChatGPT can do a better job of communicating with patients than themselves. This sentiment is echoed broadly: Earlier this year an article suggesting ChatGPT could be preferred to human doctors made a stir on social media and garnered significant media coverage (McPhillips, 2023). So, when machines, powered by AI, start defining communication goals, deciphering meaning, and suggesting how a doctor should communicate with a patient, aren't they exercising a form of agency? Our study suggests so, and while these scenarios are futuristic now, they are clearly salient in the minds of the future doctors we spoke with. Thus, in the current moment, the perceived agency of these machines isn't derived merely from their ability to communicate or take action, but from their newfound role as the timekeepers of knowledge evolution. Their rapidly improving ability to accumulate and process information is essentially redrawing the temporal boundaries of human learning, professional growth, and expertise.

Conclusion

Human-machine communication takes place in many contexts both personal and professional. Expectations, attitudes, and beliefs about machines affect the way that people interact with them. This research takes a unique study population—medical students on the cusp of long careers—to take a step backward in the timeline of attitudes toward AI. Our research shows that these attitudes can be strong even without extensive interaction with AI in the workplace. Overall, we find that attitudes are generally positive toward the use of AI, but some hesitation remains. The most salient norms are the ones medical students hope for, namely that AI is primarily a tool and acts as an ouvrier for less desirable tasks. They also believe that AI will be introduced to fulfill organizational goals, and they may not be granted autonomy to use or not use AI in these largely administrative functions. But in their personal workflows and relationships with patients, future doctors believe that they will have control over AI tools; humans remain the boss. We also witness that the introduction of a new, revolutionary technology can affect people's sense of control over their own personal development in relation to machines, and affect their confidence in a number of domains, including communication. When the future doctors we interviewed move throughout their careers, future machines will inevitably be there as well. Fortunately, for most, this future seems a better place.
Author Biographies

Andrew Prahl (PhD, University of Wisconsin-Madison) is an Assistant Professor in the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore. His research frequently compares humans and machines, including human workers make sense of machines in the workplace or society more generally understands machines. Andrew’s research on machines has spanned multiple industries including health care, aviation, journalism, humanitarian aid, and public relations.

https://orcid.org/0000-0003-3675-3007

Kevin Tong Wen Jin is a medical student in the Lee Kong Chian School of Medicine at Nanyang Technological University, Singapore. His research interests include the integration of technology into medical education and the future of AI in health care.

References


