IlLUminate Blog Transcript: Rebecca Wang on Overcoming AI Resistance in Health Care
Recorded November 4, 2024. Listen to it here.

ANNOUNCER: 00:01	[music] This podcast is brought to you by Illuminate, the Lehigh Business blog. To learn more, please visit us at business.lehigh.edu/news.
JACK CROFT: 00:11	Welcome. I'm Jack Croft, host of the Illuminate podcast for Lehigh Universities College of Business. Today is November 4th, 2024, and we're talking with Rebecca Wang about a recent study she conducted with colleagues at Lehigh University and Seattle University that suggests highlighting human bias can reduce individuals' resistance to the use of artificial intelligence, or AI, in health care. Dr. Wang holds the Class of '61 Professorship in marketing in Lehigh's College of Business. She was the lead author on the study titled "To err is human: Bias salience can help overcome resistance to medical AI," which was published recently in the journal Computers and Human Behavior. The study was co-authored by Matthew Isaac of the Albers School of Business and Economics at Seattle University; Lucy Napper, associate professor of psychology, director of undergraduate studies, and associate director of the Health, Medicine and Society Program in Lehigh's College of Arts and Sciences; and Jessecae Marsh, professor of psychology and associate dean for interdisciplinary programs and international initiatives in Lehigh's College of Arts and Sciences. Rebecca's research reflects her interest in marketing, data science, and technologies, and focuses on digital and mobile channels, social media dynamics, and data-driven marketing strategies. Rebecca, it's nice to have you back on the Illuminate podcast.
REBECCA WANG: 01:49	Thank you so much, Jack, for inviting me. It's always fun.
CROFT: 01:52	So getting to it - and let's set the stage first for the study you and your colleagues recently did - you began by reviewing prior research regarding how people generally view artificial intelligence when it's used in healthcare decision-making differently than they do in other industries, such as banking and finance, manufacturing, and retail. So what did the existing research tell you about those differences as you embarked on this study?
WANG: 02:21	Previous research on AI has been mostly focused on banking and manufacturing and retail, like you said, as those applications have been around for longer. And the implications there typically focus on efficiency and cost savings, which is important in these utility-driven function-focused types of industries. Companies streamline their processes with AI, and customers get their products and services, and everyone's happy, right? Health care, on the other hand, is much more complicated. Off the bat, it is more regulated than other industries. So for a provider to adopt and use AI health care, it needs to show that the technology does, in fact, provide greater benefits than the existing solution. And perhaps even more importantly, that it does not cause any patient harm. So from a patient's perspective, consumer's perspective, the consequence of AI health care getting it wrong could be detrimental versus, say, AI retail recommender telling me to wear the wrong colors or wrong kinds of shirts, right? That kind of consequence in comparison is minimal. So that's why health care is such a nuanced type of industry to apply AI.
WANG: 03:42	And furthermore, AI services are also very personal with respect to data sharing and the decision-making process resulting from the interactions between the providers

and the patients. So as such, health care contacts are typically very sensitive and nuanced. So for patients, individuals, right, we believe that we're all born unique. So how can an AI system account for the fact that I have this particular ailment or this particular condition, right? I mean, we seek for this type of interactions with our providers, right? And we want individualized interactions and care. And how can an AI system take care of all that? And this has been documented in previous research, which shows that patients are hesitant toward AI health care due to the perception that AI lacks this kind of nuanced understanding for people's individual cases and needs. And this kind of self-perceived uniqueness is less of an issue in other industries like manufacturing or everyday retail. But it is much more important in industries that involve complex decision-making, like health care, for instance. And this is how health care differs from AI applications in other industries.

CROFT: 05:05 It's probably helpful to start by defining some terms that are keys to your study in particular that we'll be talking about today. First is from the title and is a term that you'll be using throughout. What is bias salience, and why is it important?

WANG: 05:23 We define bias salience as, well, making bias salient, right? So making people become more aware of the potential for bias, especially in any human decision-making processes. So biases are often seen as a human and subjective flaw. And this highlights this tendency that human providers may, in fact, be more subjective and more biased in certain ways. And relatively speaking, right, AI may be more fair and less biased because robots don't judge, right, whereas humans do, right? So it's this notion of reminding people the inherent bias nature that exists in human judgments. And hopefully, by making that bias salient or reminding people that humans are inherently perhaps subconsciously biased, people become more open to accepting AI in health care. And that's where the significance lies, which is we posit, and we show that in order to reduce resistance to AI, in sensitive industries like health care, bias salience may, in fact, play a role.

CROFT: 06:51 Now, next is algorithm aversion. What is it, and how does that come into play with AI?

WANG: 06:59 Algorithmic aversion or algorithm aversion is a pretty established concept. And it refers to people's general tendency to distrust or resist using algorithms or, in other words, computing logic and functions that are embedded and automated in AI systems, even if the AI systems may actually be more accurate or reliable than humans. And this aversion is particularly noticeable in sensitive contexts like health care. As a result, patients are reluctant to accept AI-generated diagnoses or treatment recommendations, despite the potential for AI to offer cost savings, efficiency, and sometimes perhaps greater accuracy. So people perceive AI as lacking these type of nuances that human providers can offer as we discuss because we're all unique, right. And this fear is called uniqueness neglect by previous research. This perception, combined with concerns about fairness and transparency, makes people hesitant to rely on AI, particularly in health care contexts. So that's what algorithmic or algorithm aversion is. And it is a barrier to the adoption of AI in medical settings.

CROFT: 08:22 Finally-- and I think that takes us to this next term, which is AI integrity. How do you define that, and what role does it play in how people generally view medical AI?

WANG: 08:35 So we define AI integrity with two dimensions, perceived fairness and perceived trustworthiness relative to a human counterpart. So perceived fairness depends on a

provider or system's ability to make decisions that adhere to social ethics. So evaluations of fairness typically stem from comparisons, right? So if a situation results in equal treatment and outcomes for all parties, then it is deemed fair. In other words, if AI can consistently provide the accurate and the same treatment for a given set of conditions while disregarding inconsequential variables, then it is deemed a fair system. So the second dimension, perceived trustworthiness, is based on the faith and goodwill that accumulates when information is less transparent. So it's slightly different from fairness. Fairness is everyone should be equal, whereas trustworthiness is more about how much I trust the system, even though information is not transparent.

WANG: 09:46

So this is why interpretable AI or transparent AI becomes crucial in helping the patients understand and accept AI health care, so if I can't understand how AI makes a decision that increases its trustworthiness in patients' perception. By our definition, when AI integrity is perceived as high, then people are more likely to trust and accept AI recommendations as they believe the system is impartial and fair and trustworthy. Conversely, if AI integrity is questioned, maybe because of this lack of consistency or lack of transparency or this black box decision-making process, then people might resist AI health care. So, in short, we propose that AI integrity plays a mechanism type of role in influencing patients' openness to integrating AI into their health care delivery.

CROFT: 10:43

As you've mentioned, your study set out to test whether highlighting how human bias can affect health care decisions could enhance perceptions of AI integrity, thus reducing resistance to medical AI. And I'm curious what led you to pursue that question in the first place?

WANG: 11:01

I have always been fascinated by how consumers use or accept new technologies. So given that AI is up and coming and it has the potential in making health care delivery more efficient while simultaneously improving the quality of patient care, I think it's imperative for patients to be willing to consider medical AI as a potential complement to traditional care. So as we discussed, there is some previous research that has shown that people are hesitant to adopt medical recommendations and diagnoses from AI agents, right? The whole uniqueness neglect, the, "I am important, AI is not nuanced enough to make those distinctions," make those recommendations that's personalized to me by all these concerns. So there is well-documented research, highly regarded research that already talks about that or examines that. However, on the other hand, there's another stream of research that suggests there are certain circumstances under which humans might prefer machines.

WANG: 12:12

For instance, children with autism, they learn better with robots or patients preferring automated systems or AI systems over in-person help when they encounter embarrassing situations or when they need help with more of diseases with stigma kind of situations. And then there are also other research findings as well as personal anecdotes that I've come across that suggest patients might be afraid of their physicians and providers or patients want to make sure that they wouldn't offend their physicians or providers. And I also came across a few stories told by nurses, anonymously, about how they would treat patients differently depending on whether they like them or not. So all of these findings as well as personal experiences and the stories and anecdotes that I've come across led me to wonder whether algorithmic

aversion can be mitigated given all the shortcomings that human health care providers might have.

CROFT: 13:20 Cutting to the chase, with the realization that the title of the study kind of is a spoiler alert for the overall finding, what were some of the other key findings from your study?

WANG: 13:35 The findings from this paper are quite straightforward. Half of it is in the title. So, essentially, we show that bias salience reduces patients' resistance to medical AI. In other words, reminding people that biases exist and are inherent in humans' decision-making process can make them more receptive to AI in health care settings. This effect of bias salience on reducing resistance to AI was consistently found across multiple studies, across multiple variations in medical scenarios, as well as the types of biases highlighted. For instance, just general cognitive bias or gender bias or age bias. And we also show that bias salience increases perceived AI integrity. So high bias salience not only reduced patients' resistance toward AI, but it also enhances the perceived integrity of AI. So participants, in other words, view AI as more fair and more trustworthy when they're reminded that humans, in fact, are biased in nature.

CROFT: 14:51 Right. And you talked about multiple studies. And without getting too deep into the weeds about how each was conducted, I think it would be helpful to talk about the six studies you and your colleagues conducted and what they revealed about bias salience.

WANG: 15:11 Yeah. So first, we started out with a preliminary pilot study that assessed whether people associate bias more with human providers than with AI. And the results confirmed what we suspected, which is robots don't judge, right? Patients-- or participants, to be specific, in fact, perceived human providers to be more biased than AI. So this finding sets the stage for our later studies. So then we conducted four experimental studies to test whether our hypothesis, which is high bias salience, so again, in other words, making participants become more aware of human biases, could, in fact, influence their preference for medical AI versus human providers. So in two of the four studies, we simply show the treatment groups an infographic, telling them that humans' decisions can be influenced by a variety of cognitive biases, right, so recency bias or confirmation bias, for instance.

WANG: 16:21 And then in the other two studies, we ask the treatment group to reflect on gender or age bias that they may have encountered in the past. And then we ask the participants to choose between human or AI health care recommendations or whether they prefer care from an experienced nurse versus a less experienced nurse that is facilitated by an AI assistant. Across all these studies, we consistently found that high bias salience reduces resistance to medical AI. And then in our last study, we want to examine the mechanism behind this effect, focusing on AI integrity. So it basically says, "When bias salience was high, right, how do participants feel about AI system?" And just as we predicted, when you trigger people's bias salience, they think AI has greater integrity. That is, the perceived fairness and perceived trustworthiness are higher. So this essentially ties the story together, explaining the mechanism of this bias salience phenomenon that we observe. Overall, this is essentially the finding, right? So we show that we can, in fact, shift people's perceptions toward AI health care by making them aware of human biases.

CROFT: 17:52

Now, in one of the studies you were just talking about, you had asked participants to write about the reason they chose either a more experienced consulting nurse or a less experienced nurse who was assisted by an AI assistant. And one of the comments you note in the study was from a participant who chose the more experienced nurse and said, "I would rather die from human error than a bug or glitch." And that struck me as really getting to the heart of how strong AI aversion is for some people and the challenge in overcoming it.

WANG: 18:34

Yeah. Absolutely. I was chuckling when I read that. It shows that for some patients, they definitely have complete trust in human judgment, right, despite its potential fallibility. So generally speaking, from the study, we observe that there are participants like this one that prefer human providers. And they tend to-- in their reasoning, they tend to put a lot of emphasis on trust and empathy and human provider's clinical experience. On the other hand, there are participants that prefer the nurse facilitated by an AI assistant, right? So it's human plus AI. And these participants tend to focus on the information and the knowledge and the functional aspects of health care delivery. And they also prefer having a second opinion or a different option. So this is interesting, right? Not only we prime these participants with high versus low bias salience, but once they make the decision, we also ask them why. And then the reasoning-- just from reasoning, more of an exploratory type of analysis shows that patients actually have different emphasis when they search for health care-- when they acquire health care. So this particular participant who would prefer to die from human errors over a bug or a glitch basically suggests that if you want to integrate AI into health care, then you must carefully balance the technology with clear communication. And hopefully, perhaps, AI in the future can be a complement to human care, but it's never a replacement of it. And I think that message needs to be clear.

CROFT: 20:28

Yeah. And I think that brings us to the next question I have, which is about what's at stake in finding ways to reduce the resistance many patients have to the idea of accepting the use of AI in their health care? And I was struck again by a quote that was actually in your study citing a review article that was published in one of the leading peer-reviewed medical journals, Nature Medicine, that wrote, quote, "AI is poised to broadly reshape medicine in the coming years." So as AI becomes more prevalent, then, I guess we're getting to why does it really matter if people are accepting of AI or not. What's wrong if they just want human?

WANG: 21:20

Yeah. No, that's fair. I think especially there are a lot of challenges that the Nature Medicine article also highlights, which includes gathering unbiased and representative data and training the AI systems fairly, given that the data may not be representative. So there are definitely challenges. And I think people's algorithmic aversion toward AI medicine on some level could be justified. This said, though, the potential benefits are also profound, such as using data for diagnostics and personalizing treatment options. It also streamlines the health care delivery process, right, perhaps making wait times shorter or allowing health care to be more equitable or more accessible to everyone. So for instance, virtual health care can make health care more accessible.

WANG: 22:16

And it can also potentially even help with early disease detection or even drug discovery. So instead of just doing annual checkups and everything seems OK, you can actually kind of predict what your next set of numbers might look like, right, as someone ages. And then, hopefully, you can catch or predict. Even though the

numbers are OK now, but maybe five years down the road, right, this patient might encounter X, Y, Z problems. So all of this, in theory, could be done and I'm pretty sure in some circles are being done already. And hopefully, this will broaden out. But for this to broaden out, patients need to be willing to try it, right, to be willing to be open to AI medicine. So I think these are the stakes as we try to find ways to reduce AI resistance.

CROFT: 23:19

And your study also toward the conclusion actually mentions what I thought was a really interesting question for future study. But it's the ethical implications for health care providers and policymakers in deliberately highlighting the prospect of discriminatory human bias as a way to increase acceptance of AI. In other words, reminding people up front that there are biases in health care, which is beyond discussion, I think, for people who have seen the literature, but it is not the most pleasant thing for most people to think about that. Because of their race, their gender, their age, they may be getting a lesser degree of health care than some other people might be, and that AI is the way to rectify that. So I know this is outside the scope of your study because you raised it to say this is something that should be further studied. But I do wonder if you have any thoughts on how that might be accomplished.

WANG: 24:35

Yeah, this is a great question. And I don't really have an answer, but I do want to point out that a lot of these biases or making biases salient, I don't think it takes that much to remind patients that humans are inherently biased. And it's not that you're getting less care, I don't think, but perhaps a more biased care. So for instance, a patient who is obese walk into a clinic, then immediately, right, it's possible that a physician or a provider would think about more diabetic or diabetes-related type of diseases because historically, that's typically what a diagnosis is. And AI system, as of now, probably would make the same kind of judgment as well, right?

WANG: 25:28

So it's a work in progress, I think, AI systems, to be sure. But hopefully, it will be a useful tool to bring forth better diagnosis and targeted treatment. So all this said, I just want to say that there aren't-- I don't think it takes that much to remind patients that humans are inherently biasing. In our studies, all we did was we show participants' infographics that describe different types of cognitive biases. And then we also ask them to think about their past experiences, whether they encounter biases, or we cited research findings on gender-related biases and whatnot in health care communications and recommendations. So all these subtle reminders, I think, would make patients become more cognizant.

WANG: 26:19

But I don’t think that patients need to be too worried. I think knowledge is power, right? So if you know these things going in, then perhaps you are more open, and hopefully, you're capable of asking more relevant questions of your providers. And then there's also-- from the provider standpoint, I think transparent communication is key. I think solving this might be more important from a provider standpoint, which is providers and policymakers should clearly communicate the presence and impact of human biases in their decision-making. So providers should admit that their own preexisting biases-- maybe admit to themselves, too. It's not necessary to the patients, but at least to themselves that their own preexisting biases may be subconsciously affecting their judgments. So having this AI system as an assistant can help them, and not replace them, to make better decisions, especially in certain situations. Perhaps not in all situations, but in certain types of situations. And then

policymakers can launch educational campaigns that explain how biases can affect diagnoses and treatments, perhaps using real-world cases or examples to illustrate how AI can improve the care delivery process. It's very sensitive, right, highlighting human bias, but I think it can be done with care and transparency and hopefully with a balanced narrative that can showcase how humans and AI can complement each other without diminishing the value of human providers.

CROFT: 28:09 In the final section of your study titled Future Directions of Research and AI Human Interaction, I thought there was something really interesting and, again, that struck me that it kind of gets at, I think, at least some of, if not a lot, of what we've been talking about here today. You and your colleagues write - and I'll quote - "The challenge of medical AI acceptance is not only technological but also psychological, rooted in the complex and deeply personal nature of health care and medical interactions." I'm wondering if you have any closing thoughts on that challenge.

WANG: 28:52 Addressing this challenge involves more than just providing or proving that AI works, right? It requires fostering trust in implementing a clear communication strategy between the providers and the patients. It requires the health care providers admitting that as humans, they make mistakes where they have biases and AI is there to help to mitigate those mistakes and those biases. It requires health care providers to emphasize with their patients that AI is a complementary tool rather than a replacement. And together, they hopefully can provide better care for the patient while accounting for all the nuances and the uniqueness that each patient's case can present.

WANG: 29:41 And then I think both patients and health care providers should take care to learn about how AI works and its pros and cons. And patients often seek empathy and the sense of wanting to be understood, particularly in health care situations where things could be potentially complex and individualized. And these are the qualities that are traditionally associated with human providers. And we want to make sure that patients understand that these qualities will continue even as AI is integrated as part of the delivery chain. And I also want to add that having an AI system in place might also remind the providers to keep their biases in check. Sometimes a lot of these biases are subconscious. But knowing that there is an agent working alongside with you-- an AI agent working alongside with you, you might then question yourself, right, that, "Is my own personal inherent biases influencing the way I make decisions?" So all of these, I think, can work in conjunction together to help deliver a better care for the patients.

CROFT: 30:59 Rebecca, I want to thank you again for being with us today. I told you this before we went on, I mean, I really found this study interesting and insightful in a lot of ways in terms of what's happening in a lot of different fields, but particularly the role of artificial intelligence in the medical field in particularly does, I think, feel different to most people than it does with the financial app or in a lot of other business fields. So glad that you're here looking at this and pursuing it.

WANG: 31:37 Yeah. Yeah. Thank you so much. Yeah. These are wonderful questions. So it helps me think, too, about next steps.

CROFT: 31:45 I'd like to, once again, thank my guest, Rebecca Wang. Her research has been published in such leading journals as Journal of Marketing Research, Journal of Consumer Research, Journal of Retailing, Journal of Interactive Marketing, and

Computers and Human Behavior. This podcast is brought to you by Illuminate, the Lehigh Business blog. To hear more podcasts featuring Lehigh Business thought leaders or to follow us on social media, please visit us at business.lehigh.edu/news. This is Jack Croft, host of the Illuminate podcast. [music] Thanks for listening.