Rationally Trust, but Emotionally?
The Roles of Cognitive and Affective Trust in Laypeople's Acceptance of AI for Preventive Care Operations

Accepted at *Production and Operations Management*

Nakyung Kyung
National University of Singapore

Hyeokkoo Eric Kwon
Nanyang Technological University, Singapore
eric.kwon@ntu.edu.sg

Abstract

Artificial intelligence (AI) is transforming healthcare operations. Nevertheless, particularly in the context of preventive care, little is known about how laypeople perceive and accept AI and change their behavior accordingly. Grounded in a solid theoretical framework of trust, this study bridges this gap by exploring individuals' acceptance of AI-based preventive health interventions and following health behavior change, which is critical for preventive care providers' operational and business performance. Through a randomized field experiment with 15,000 users of a mobile health app complemented by a survey, we first show that the use and disclose of AI in preventive health interventions improve their effectiveness. However, individuals are less likely to accept and achieve health behavior change suggested by AI than when they receive similar interventions from health experts. We also observe that the effectiveness of AI-based interventions can be improved by combining them with human expert opinions, increasing their algorithmic transparency, or emphasizing their genuine care and warmth. These results collectively suggest that, different from conventional technologies, AI's deficient affective trust, rather than comparable cognitive trust, plays a decisive role in the acceptance of AI-based preventive health interventions. This study sheds light on the literature on the role of new-age technologies in behavioral operations management and consumer marketing as well as the role of trust in technology acceptance. Valuable practical implications for more effective management of AI for preventive care operations and promotion of consumers' health behavior are also provided.

Keywords: Artificial Intelligence (AI), Technology Acceptance, Affective and Cognitive Trust, mHealth (mobile health) app, Randomized Field Experiment

1. Introduction

The world is experiencing a serious shortage of human resources for healthcare operations. The shortage of healthcare workers was estimated to be 17.4 million in 2013, and it is projected to decline by only 17% by 2030 (World Health Organization 2016). With the growing emphasis on and corresponding escalating demand for *preventive care* in particular, the shortage of its human resources has become especially severe (Stein and Brooks 2017). Physicians suffer from a lack of time and resources for preventive care, such as prescriptions for diet and exercise (Douglas et al. 2006). As a result, people who could benefit from intensive in-person behavioral counseling often receive inadequate services (Stein and Brooks 2017). Specifically, while 75% of the U.S. healthcare budget is allocated to preventive care (Beaton 2017), only 22.4% of Americans receive recommended services (Borsky et al. 2018).

Against this backdrop, by addressing the gap between the high demand for and low supply of preventive care, artificial intelligence (AI) has drawn considerable attention as a transformative force to preventive care operations (Guha and Kumar 2018, Hopp et al. 2018). Specifically, AI is shown to improve preventive care operations by providing more cost-efficient and personalized *preventive health interventions*, including dietary recommendations (Kong and Tan 2012) and exercise prescriptions (Zhou et al. 2018), which have been offered by health experts. Accordingly, AI is transforming how patients interact with health experts (e.g., physicians) for preventive care.

However, AI-based preventive health interventions (hereafter *AI-interventions*) are facing a critical challenge: users' significantly lower trust in AI for preventive care operations compared with their trust in health experts. For example, in one survey of 2,048 U.S. adults, only 20% of individuals indicated that they would trust AI which generates healthcare advice (Schierber 2019). Given that trust is a key challenge regarding users' acceptance of AI applications (Kumar 2021), such low trust can hamper the effectiveness of AI-interventions. This creates complexities for preventive care providers regarding whether and how to adopt such a new-age technology to replace their existing preventive health interventions provided by health experts (hereafter *Human-interventions*).

As the performance of AI has been dramatically improved in recent years, firms actively promote that their AI applications provide comparable or even better performance at a lower cost compared to human counterparts. Given its nature, the media also focuses on such performance comparison, which often includes exaggerated claims about AI's performance (Oscar 2018). Accordingly, laypeople have heightened their expectations about AI's performance (Brynjolfsson and Mcafee 2017), and the media hype surrounding several major events where AI beats or outperforms humans (e.g., AlphaGo defeated the world champion of Go) amplify such perception. For example, a recent report showed that more than 66% of patients expect better performance from healthcare AI than from health experts (Collier et al. 2017).

Therefore, based on the general rationality-based trust in conventional technologies, which is referred to as cognitive trust (Komiak and Benbasat 2006), it is surprising that laypeople have such low trust in AI for preventive care operations despite their growing expectation regarding its performance. This is because cognitive trust is established and developed primarily by observing objective performance (Lewis and Weigert 1985).

To identify the source of this discrepancy, given that AI is one of the new-age technologies with distinct features from conventional technologies (Kumar 2021), we revisit the extant technology acceptance theories based on an in-depth theoretical discussion about the unique characteristics of AI. In short, different from a general technology of which acceptance is often determined based on its own features or in comparison with those of other comparable *technologies*, individuals tend to first compare an AI application directly to its *human* counterparts (Gursoy et al. 2019). Such a comparison calls for expanding the scope of trust in the previous technology acceptance studies from cognitive trust into affective trust. Affective trust encompasses emotional and somewhat irrational feelings which are not necessarily based on objective performance (Komiak and Benbasat 2006). In other words, based on their observation of its objective performance, laypeople would have comparable or sometimes even higher cognitive trust in AI than human counterparts. However, at the same time, they would have inherently lower affective trust in AI compared to its human counterparts. To lend empirical support for the discrepancy between affective and cognitive trust, we conducted a field survey and measured individuals' trust in AI and health experts at a granular level. The survey results also support that individuals exhibit *comparable* cognitive trust in AI and health experts; however, their affective trust in AI is significantly *deficient* compared with that in health experts.

Given that affective trust does not necessarily move together with cognitive trust, a more important practical question would be the relative importance of affective and cognitive trust in laypeople's acceptance of AI for preventive care operations: If affective (cognitive) trust plays a decisive role, individuals' acceptance of AI-interventions will be lower than (comparable to) that of Human-interventions, which would limit (promote) the positive impact of AI on preventive care operations. Thus, the goals of this study are to (1) assess the relative importance of affective and cognitive trust in individuals' acceptance of AI for preventive care operations, (2) thereby examine whether AI can effectively replace health experts in preventive care operations, and (3) provide theory-driven, practically validated strategies to further improve the decisive trust (i.e., affective or cognitive trust) in AI and consequent effectiveness of AI for preventive care operations.

To achieve these goals, we conducted a randomized field experiment with 15,000 unique users of one of the most popular mobile health apps in South Korea. Specifically, we provided two treatment groups with AI- or Human-interventions in which AI or health experts recommended a personalized step goal to users, respectively. A control group was constructed by designing Neutral-interventions that provided the same goals without informing the source of the goal generation (i.e., AI or health experts). We compared the effectiveness of AI-, Human-, and Neutral-interventions in terms of users' acceptance and following actual health behavior change. Our key finding is that, contrary to the recent findings on the negative impact of AI disclosure (e.g., Luo et al. 2019), AI-interventions exhibit significantly higher effectiveness compared with Neutral-interventions. However, they are significantly less effective than Human-interventions. This implies that affective trust in which AI is deficient, rather than comparable cognitive trust, would be particularly important in individuals' acceptance of AI for preventive care operations. Specifically, in the context of preventive care, the ultimate adoption decision is made by laypeople who usually have limited knowledge to make objective assessments of preventive health services. Therefore, it is difficult for them to build knowledge-driven, cognitive trust in healthcare providers. Thus, especially with regard to AI for preventive care operations, this study highlights the important role of affective trust, which has been considered less relevant to the acceptance of general technologies (Gefen et al. 2003).

In addition, we extended our survey and experiment to examine whether the enhancement in deficient affective trust in AI would in turn improve the effectiveness of AI-interventions. To improve affective trust, grounded in the theories of affective trust, we revealed the use of AI in tandem with health experts, disclosed more transparent information about the AI algorithm, or emphasizing its genuine care and warmth. The additional survey confirms that such features actually improve individuals' affective trust in AI. In addition, the results of the additional experiments demonstrate that such an improvement in affective trust in AI results in greater effectiveness of AI-interventions. These results provide further support for the decisive role of affective trust in AI for preventive care operations as well as fruitful practical guidance on how to manage AI-based preventive health interventions more effectively by improving their affective trust.

By carefully examining how users perceive and accept AI-interventions, this study contributes to the growing body of literature on the role of new-age technologies in behavioral operations management (Donohue et al. 2020). Specifically, we focus on the management of AI for preventive care operations, thereby responding to the recent call for research on how AI reshapes and transforms healthcare operations (Guha and Kumar 2018, Johnson et al. 2020). In addition, this study offers novel theoretical insights into the role of affective trust in users' acceptance of AI, thereby making significant contributions to the theories of technology acceptance and trust for more efficient operations management (Ha et al. 2011). Moreover, this study also provides fruitful managerial implications for the design of more effective AI-interventions, which is a key to the success of preventive care business, thereby contributing to research on the use of AI for consumer marketing in general and promotion of consumers' health behavior in particular (e.g., Kumar 2021).

2. Research Background

AI refers to machines performing *human-like, cognitive* functions, such as perceiving, reasoning, learning, problem-solving, and decision-making (Rai et al. 2019). During the past few years, substantial resources have been allocated to a wide range of AI-related products, services, and business operations, including virtual assistants (e.g., Alexa and Siri), cashier-less stores (e.g., Amazon Go), and supply chain and inventory management (Ellis et al. 2018). Accordingly, the global business value of AI is expected to reach $3.9 trillion in 2022 (Gartner 2018).

Among others, AI is revolutionizing the healthcare operations in particular (Guha and Kumar 2018, Johnson et al. 2020). Healthcare can be generally categorized into two types: diagnostic (i.e., investigating and treating specific health issues) and preventive care (i.e., screening and preventing such health issues). The transformative power of AI is equally promising in both areas. For diagnostic care, due to their powerful predictive performance, AI-based applications assist healthcare providers in making better decisions in various contexts, such as cancer prognosis (Johnson et al. 2020) and intensive care monitoring (Zhang and Szolovits 2008). For preventive care, with a goal to improve users' health behavior, a number of mobile apps are providing AI-based preventive health services and interventions, such as suggesting personalized exercise goals (Okano et al. 2013) and helping people estimate and monitor their calorie consumption in real-time (Pouladzadeh et al. 2014). Extensive studies on healthcare AI have examined how to improve its *technical* performance, including accuracy, cost-efficiency, and scalability (Camerer et al. 2019, Johnson et al. 2020). Even if healthcare AI achieves exceptional technical performance, however, it cannot improve healthcare providers' *operational* and *business* performance unless its end-users (e.g., physicians in diagnostic care, patients in preventive care) accept and use it and change their behavior accordingly. Thus, end-users' perceptions and actual behavior regarding healthcare AI are critical. Nevertheless, relatively scant attention has been paid to such areas.

While recent studies in operations management (e.g., Cui et al. 2021, Fan et al. 2020) and marketing (e.g., Castelo et al. 2019, Logg et al. 2019, Longoni et al. 2019, Luo et al. 2019) have investigated how individuals perceive, accept, and use algorithm- (e.g., AI-) based services in diverse industries, this study contributes to five notable research gaps in the literature. First, while several studies have focused on AI for diagnostic care (e.g., Fan et al. 2020, Longoni et al. 2019), this study is among the first to investigate laypeople's perceptions and behavior with respect to AI-based services in the context of preventive care, which accounts for more than 75% of U.S. healthcare spending (Beaton 2017).

Second, in the context of preventive care, this study provides a novel insight into which theoretical construct drives individuals' acceptance decision and actual behavior change regarding AI-based services. This is particularly important to understand the complexities regarding whether and how to adopt this new-age technology (Kumar et al. 2021). While Longoni et al. (2019) and Fan et al. (2020) provide useful insight in the context of diagnostic care, their results can hardly be generalized into AI for preventive care. Specifically, Fan et al. (2020) focused on acceptance of AI-based diagnosis support systems by healthcare professionals, who have extensive medical knowledge and might thus behave differently toward AI from laypeople. Similarly, Longoni et al. (2019) identified that uniqueness neglect, a concern that AI is less able than humans to consider consumers' unique characteristics, drives consumers' different willingness to use AI- and human-based services in the context of diagnostic care. However, based on our additional follow-up survey, we identified that uniqueness neglect might not be a significant predictor of individuals' acceptance of AI-based services in our research context of *preventive care* (see Online Appendix for more details). In this study, different from conventional technologies of which acceptance decision would be governed by cognitive trust, we identified that affective trust is a decisive factor in individuals' acceptance of AI for preventive care operations, contributing significantly to the technology acceptance and trust theories.

Third, given the limited understanding of which theoretical construct drives individuals' acceptance of AI for preventive care operations, it remains relatively silent in terms of developing theory-driven, practically validated strategies to improve its effectiveness. Against this backdrop, grounded in the theories of affective trust, we proposed and empirically validated that combining AI-interventions with human expert opinions, increasing their algorithmic transparency, or highlighting their genuine care and warmth improves individuals' affective trust and acceptance behavior toward AI for preventive care operations. Accordingly, we respond to the call for research on the role of trust in behavioral operations (Donohue et al. 2020).

Next, while much insight has been gained into how AI affects individuals' behavioral intention or perception (e.g., Castelo et al. 2019, Fan et al. 2020, Logg et al. 2019, Longoni et al. 2019), relatively scant attention has been paid to whether and how AI consequently influences their actual behavior. This is mainly because of the difficulty in gathering relevant information on actual behavior. Against this backdrop, drawing on advanced mobile technologies (e.g., mobile apps, smartphone pedometers), this study contributes to the literature on behavioral operations management (e.g., Donohue et al. 2020) and consumer marketing (e.g., Kumar 2021) by examining the impact of AI on individuals' actual behavior in a preventive care context. Specifically, our outcome variables measured in the field, i.e., individuals' acceptance of AI and consequent health behavior change, have important implications for public health as well as preventive care providers' operational performance and marketing effectiveness. In particular, this study belongs to the "improve behavior" category (e.g., Choudhary et al. 2021), which has been outlined as a key research goal of behavioral operations management (Donohue et al. 2020), and thereby contributes to the operations management literature that employs behavioral interventions to improve individual behavior and consequent operational performance (e.g., Bendoly 2013, Kim et al. 2020, Song et al. 2018).

Lastly, while most extant studies with surveys and lab experiments (e.g., Castelo et al. 2019, Fan et al. 2020, Logg et al. 2019, Longoni et al. 2019) could mimic actual situations, their results would be vulnerable to observer effects. Thus, they might be affected by a lack of external validity and biased estimates (Aral and Walker 2011). In this study, a randomized field experiment, which has been recognized as an important method for behavioral operations management research (Ibanez and Staats 2018, Nguyen and Kim 2019), allows us to obtain a population of users with the real motivations that drive health behavior (Baek and Shore 2020) and thereby less biased and more generalizable estimates of causal effects.

3. Hypotheses Development

With the introduction of the Internet, traditional offline relationships have moved to online platforms where individuals perceive greater risk in their relationships (Reichheld and Sasser 1990). Accordingly, trust has emerged as a relatively new but important dimension of technology adoption (Gefen et al. 2003). Trust reduces risk perceptions and uncertainty regarding the utility of a technology when the utility is not immediately verifiable, thereby increasing its acceptance (Gefen and Straub 2004). The existing technology acceptance studies emphasize the rationality of users regarding the adoption of a new technology and assume that users make their adoption decision by carefully assessing the expected utility of the technology (Venkatesh et al. 2016). Thus, trust has usually been conceptualized as a deliberate and *rational assessment* of a trustee’s characteristics that trustors rely upon (Komiak and Benbasat 2006). This rationality-based trust is specifically referred to as cognitive trust. Cognitive trust is deeply rooted in the rational expectations that a trustee will bring utility and advantage (Komiak and Benbasat 2006, Özer et al. 2014). Prior studies suggested that cognitive trust is accumulated by objective outcomes such as better performance (Lewis and Weigert 1985, McAllister 1995). Cognitive trust has been shown to increase acceptance and corresponding use of a technology (Komiak and Benbasat 2006); a technology with low cognitive trust is less preferred and accepted, even if careful implementation efforts are devoted (Pi et al. 2012).

As technology advances, AI has shown astonishing performance, attested by decades of research showing that statistical models and algorithms generally outperform human intuition in diverse operational situations (e.g., Preil and Krapp 2021). Healthcare AI research has also shown that AI-based health services and interventions provide more timely results with lower error rates than humans in diagnosing complex diseases (Johnson et al. 2020) as well as in generating user-specific goals for daily exercise and food intake (Kong and Tan 2012, Zhou et al. 2018). In addition, it has been identified that even a simple linear regression model outperforms human experts in diagnosing medical and psychological illnesses (Dawes 1979, Grove et al. 2000). This outperformance of AI over humans has also been gradually acknowledged by laypeople. An industry report shows that more than 66% of patients agreed that healthcare AI had better performance (e.g., AI could assess greater amounts of data and provide a more accurate prediction for diseases) than human experts (Collier et al. 2017). Accordingly, individuals' cognitive trust in AI would be comparable to or sometimes even higher than their cognitive trust in human counterparts.[1]

Trust can also encompass emotional feelings related to assurance and comfort, such as kindness, caring, bonding, and openness, which is referred to as affective trust (Ha et al. 2006, Komiak and Benbasat 2006). Affective trust has received relatively little attention in the technology acceptance research, as it was considered "arguably irrelevant to a business transaction" (Gefen et al. 2003, p. 60). Moreover, because affective trust is far less commonly formed with objects than humans (LaRosa and Danks 2018), this aspect could have been safely neglected in the previous acceptance research on conventional technologies. A general positive correlation between affective and cognitive trust also contributes to this tendency. Specifically, on the one hand, cognitive trust increases confidence in the utility and usefulness of a technology (Lewis and Weigert 1985), which provides a base for emotional bonding toward the technology and thereby improves its affective trust (Johnson and Grayson 2005). On the other hand, affective trust in a technology can also heighten its cognitive trust (Punyatoya 2019) because, to a certain extent, affective trust can act as emotional security that ensures the receipt of expected benefits from the use of the technology (Rempel et al. 1985). This reinforcing positive correlation between cognitive and affective trust has been examined in various contexts, including the adoption of recommendation agents (Nicolaou and McKnight 2006), online financial services (Pi et al. 2012), and e-commerce platforms (Punyatoya 2019).

However, such emotional affective trust is not necessarily positively correlated with cognitive trust (Cummings and Bromiley 1996, McAllister 1995) and can even be perceived as irrational on *some* occasions (Gefen et al. 2003). Compared with general technologies, AI is unique in several ways, and the most notable difference is that it imitates *human* intelligence. In other words, most AI applications are developed to replace existing human tasks. Thus, individuals tend first to compare AI with human counterparts when determining whether to accept and use it (Gursoy et al. 2019). Such a comparison between a technology and its human counterpart makes individuals consider not only cognitive trust but also affective aspects of the technology.

The human-machine interaction (HMI) literature has demonstrated that while machines, including AI, often outperform humans in terms of their cognitive capabilities, they usually lag far behind humans in terms of affective characteristics (Haslam et al. 2008). This is because individuals deem machines incapable

---

[1] Our field survey also provides empirical support for the arguments regarding individuals' relative trust in AI and health experts. Specifically, the survey results demonstrate that individuals have statistically similar cognitive trust toward AI and health experts. On the other hand, individuals exhibit significantly lower affective trust in AI than health experts. The details of the survey are provided in Section 4.

of experiencing emotion and sensation, given that they are designed to perform cognitive tasks in a standardized and rote manner (Haslam et al. 2008, Turkle 2005). Therefore, while AI might gain comparable or sometimes greater cognitive trust from individuals than its human counterpart, it would be deficient in affective trust: individuals' affective trust in AI would be lower than their affective trust in health experts.[1] Such low affective trust in AI in preventive care settings has been demonstrated in the literature. For example, a study about disease screening, a representative preventive setting (U.S. Preventive Services Task Force 1996), found that users generally perceive less affective trust toward AI compared with a human expert (Ongena et al. 2020). Specifically, users generally agreed that AI does not take their feeling into account and that humans are more responsible compared to AI. Another study on individuals' usage intention and perceived risks regarding AI-based skin cancer screening found that increased patient anxiety was its most commonly perceived risk compared to clinician-based procedure (Nelson et al. 2020). In addition, Yokotani et al. (2018) compared individuals' trust in AI and human experts proceeding mental health interviews, a common procedure for depression screening. Their findings show that participants are less likely to perceive trust and emotional rapport with respect to AI than human experts.

Particularly in a healthcare context, compared to cognitive trust, the role of affective trust in individuals' acceptance decision would be more salient, mainly because the relationship between healthcare providers and consumers is governed by significant information asymmetry. In other words, individuals often lack the ability to make objective assessments of healthcare services they receive (Alford and Sherrell 1996). This information asymmetry makes it difficult for them to build knowledge-driven, cognitive trust in healthcare providers. Accordingly, instead of cognitive trust which requires objective observations and evidence, affective trust usually works as a major indicator of service quality in healthcare. That is, rather than logically and rationally assessing available observations and evidence, individuals tend to use affective information, such as healthcare providers' human aspects, soft skills, and social and cultural backgrounds, as a determinant of their acceptance decision (Halpern 2003). Such a tendency would become even more prominent in the context of preventive care (compared to diagnostic care) where the effectiveness of preventive treatments can hardly be measured accurately due to a number of confounders, and available rough measures of its effectiveness also require a longer time.

In sum, the crucial role of affective trust particularly in a preventive care context, together with individuals' lower affective trust in AI than health experts in general, would result in their lower acceptance of AI-interventions than Human-interventions. Hence, we posit the following hypothesis:

Hypothesis 1*: Individuals will exhibit lower acceptance of AI-interventions compared with Human-interventions.*

The acceptance of products or services directly reflects a *positive behavioral intention* to use them or readiness to perform a given behavior (Schifter and Ajzen 1985). Behavioral intention represents "instructions that people give to themselves to behave in certain way" (Triandis 1979). Specifically, intentions encompass both the direction (e.g., to do or not to do) and the intensity (e.g., how much time and effort a person is prepared to expend in order to do) of a decision (Sheeran 2002). Thus, a positive or negative intention of an individual indicates that she already determined the direction and intensity of her decision and is likely to actually behave in that way. In other words, if an individual has a positive intention (e.g., the acceptance of products or services), this implies that she is ready to spend time and effort to perform and engage in a given behavior (Schifter and Ajzen 1985). Accordingly, positive intentions have been used to predict a wide range of actual behavior changes, including diet (e.g., McCoy et al. 2017), physical activity (e.g., Sheeran and Orbell 1999), weight loss (e.g., Kreuzfeld et al. 2013), and smoking cessation (e.g., Norman et al. 1999). Particularly for technology (e.g., AI), the positive link between behavioral intention to use it and its actual usage has been established in the extensive literature (see Davis 1989, Venkatesh et al. 2003).

Indeed, previous studies have shown that the acceptance of health interventions or recommendations lead to actual health behavior change. For example, Hurling et al. (2007) identified that individuals who accepted a physical activity program were more likely to lose weight than those who did not accept it. McCoy et al. (2017) provided more direct evidence for the positive link between individuals' acceptance of interventions and their actual behavior change in the context of preventive care. In this study, participants who accepted preventive health interventions spent more time walking and running compared with those who did not accept the interventions. Similarly, Kreuzfeld et al. (2013) showed that people who accepted a voluntary physical activity program for six months reduced body fat significantly more than those who did not accept the program. Thus, we expect that individuals exhibit consistent behavior toward their acceptance of different interventions and actual health behavior change after accepting the interventions:

Hypothesis 2*: AI-interventions induce less health behavior change compared with Human-interventions.*

Hypotheses 1 and 2 predict the lower effectiveness (i.e., acceptance and health behavior change) of AI-interventions than Human-interventions. If the lower effectiveness of AI-interventions is attributable to the lack of affective trust, enhancing deficient aspects of affective trust in AI-interventions would significantly improve their effectiveness. Affective trust consists of two theoretical dimensions, i.e., benevolence and integrity (Doney and Cannon 1997, Morgan and Hunt 1994), which refers to individuals' perception about how much the other party cares about them or personal attachment to other agents and that about the other party's good faith and honesty, respectively. On the one hand, previous studies suggest a positive relationship between benevolence of a service and its human aspects. This association is deeply rooted in the nature of benevolence; benevolence is formed by *personal* traits such as a warm, kind, and caring attitude as well as their delivery to another agent (Hoejmose et al. 2012, Mayer et al. 1995). Machines are thus generally considered agents lacking in benevolence (Martelaro et al. 2016). Therefore, enhancing human aspects of a machine can improve its benevolence (e.g., Tapus et al. 2007). In this regard, AI-interventions used in tandem with human experts will improve their effectiveness, as the presence of health experts would compensate for the lack of human aspects and corresponding benevolence of AI-interventions.

On the other hand, integrity is positively linked with the transparency of a service (Mayer et al. 1995). Providing transparent information on the process, criteria, and constraints of an agent's decision-making creates the perception that the agent takes responsibility for its decisions, thereby conveying trust in general (Buell et al. 2021, Özer et al. 2014) and honesty and integrity in particular (Mayer et al. 1995) toward the agent. For example, Gatling et al. (2017) show that transparent communication by a healthcare provider increases the perception that the provider will behave with high integrity. This is because transparent communications make it easier to understand how a service will be provided, thereby improving word-deed alignment perception. Thus, with improvement in integrity, AI-interventions will show greater effectiveness if more transparent information on how AI generates AI-interventions is provided. Hence, we expect that AI-interventions would be more effective when complemented by human expert opinions or algorithmic transparency:

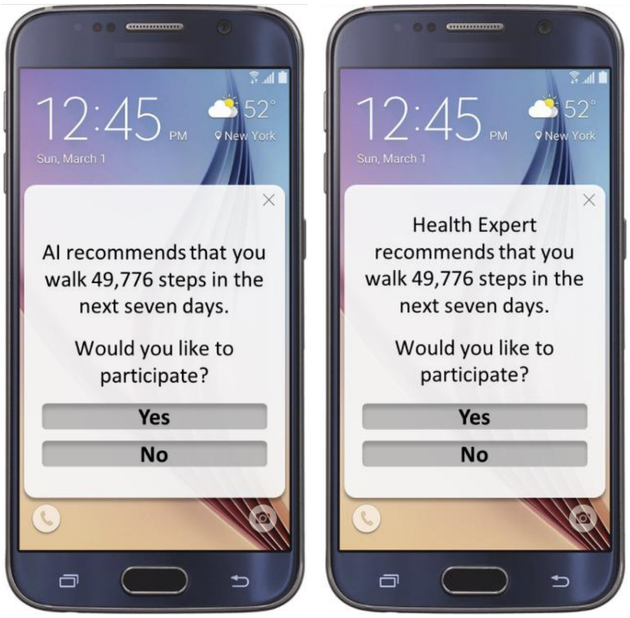
Hypothesis 3*: AI-interventions will be more effective if they are used in tandem with human experts.*

Hypothesis 4*: AI-interventions will be more effective if transparent information on their generating mechanism is provided.*

4. Hypotheses Validation

4.1. Relative Effectiveness of AI-interventions and Human-Interventions

We provided the theoretical arguments and a series of anecdotal evidence that average individuals have comparable (or higher) cognitive trust and lower affective trust in AI than human experts. Given this discrepancy between cognitive and affective trust, together with the crucial role of affective trust in preventive care, we hypothesized lower effectiveness (i.e., acceptance and health behavior change) of AI-interventions than Human-interventions (i.e., Hypotheses 1 and 2, respectively). To validate the hypotheses and provide empirical support for its underlying mechanism, we conducted both a survey and a randomized field experiment.



AI recommends that you walk 49,776 steps in the next seven days.

Would you like to participate?

Health Expert recommends that you walk 49,776 steps in the next seven days.

Would you like to participate?

Figure 1. Description of Intervention Messages for AI-and Human-Interventions Group

We first conducted the survey to validate (1) individuals' greater acceptance of AI-interventions than Human-interventions (i.e., Hypothesis 1) and to support its underlying mechanism, i.e., (2) individuals have comparable (or greater) cognitive trust in AI than health experts, (3) while exhibiting lower affective trust in AI than health experts. To this end, we designed AI- and Human-interventions which recommend individuals to walk more than a specific number of steps in the next seven days. In other words, our interventions were designed to promote individuals' physical activity, which has been considered a representative preventive care context in national guidelines for preventive care (e.g., U.S. Preventive Services Task Force 1996) as well as in the literature on preventive care (e.g., Sallis 2015). Each of AI- and Human-interventions disclosed whether the step goal was generated by AI or health experts, respectively. The participants were randomly assigned to AI- or Human-interventions. The survey started with providing each participant with a mobile screenshot of her assigned intervention. Specifically, Figure 1 shows the two screenshots shown to participants who were assigned to the AI- (i.e., the left side of Figure 1) and Human-interventions (i.e., the right side of Figure 1) group.[2] We then measured users' acceptance of AI- and Human-interventions by asking them which button they would tap in the provided intervention: Yes or No. Next, we measured their cognitive and affective trust toward an intervention provider (i.e., AI and human experts). Trust is measured using the scale developed by Gefen (2002) in three dimensions, i.e., ability, benevolence, and integrity. Ability, which indicates individual perceptions about an agent's knowledge and competence in providing higher-quality health recommendations, belongs to cognitive trust. On the other

---

[2] The step goal used in the survey (i.e., 49,776 steps) is the average value of step goals generated by an AI algorithm adopted for our field experiment. The details of the algorithm and the experiment are provided later in the manuscript.

hand, benevolence and integrity are considered to constitute affective trust (Doney and Cannon 1997, Morgan and Hunt 1994). These three dimensions of trust allow us to identify which type of trust (i.e., cognitive or affective trust) is critical for individuals' acceptance of AI-interventions. All items were rated using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).[3] The survey questionnaires were distributed at five offline community health centers in South Korea. We collected 95 valid responses while discarding three low-quality samples.[4]

Using the survey, we first examined the acceptance rates of each intervention group. As can be seen in the third column of Table 1, the Human-interventions group shows a significantly greater acceptance rate than the AI-interventions group (*t*-test, $p < 0.05$), which supports Hypothesis 1. We then compared the mean values of the three dimensions of trust (i.e., ability, benevolence, and integrity) between the AI- and Human-interventions groups (see the last three columns of Table 1). The results indicate that the AI-interventions group shows statistically similar cognitive trust (i.e., ability) in the intervention provider compared to the Human-interventions group (*t*-test, $p > 0.1$). However, survey participants showed significantly low perceived benevolence and integrity (i.e., affective trust) regarding AI than health experts (*t*-test, $p < 0.01$). This implies that, consistent with our theoretical arguments, individuals' lower acceptance of AI-interventions than Human-interventions can be attributed to their lower affective trust, rather than comparable cognitive trust, in AI. In other words, affective trust plays a decisive role in individuals' acceptance decision of AI-interventions.

Table 1. Main Survey Results

Group	Obs.	Acceptance (%)	Ability	Benevolence	Integrity
			(Cognitive Trust)	(Affective Trust)	
AI-interventions	49	34.7	2.673	2.484	2.479
Human-interventions	46	52.2	2.695	3.722	3.233
Difference (Human - AI)		17.5	0.022	1.238	0.754
p-value of *t*-test		$p < 0.05$	$p > 0.1$	$p < 0.01$	$p < 0.01$

In addition to the survey, the randomized field experiment was conducted to confirm the validity of the greater acceptance of AI-interventions than Human-interventions (i.e., Hypothesis 1) in practice. Moreover, it allows us to further investigate whether individuals exhibit consistent behavior regarding AI- and Human-interventions in terms of their actual health behavior change after accepting the interventions (i.e., Hypothesis 2). To this end, we collaborated with one of the largest mHealth app providers in South Korea. The focal app is a free app with an average of 120,000 weekly active users and has been downloaded more than 500,000 times (as of January 2021). The app runs in the background to track each user's real-time

---

[3] The survey items for the three theoretical constructs, i.e., ability, benevolence, and integrity, with the supporting literature are provided in Table A1 in Online Appendix.

[4] The quality of the survey is described in detail in Online Appendix.

walking activity using sensors in the smartphone unless users force-stop the background function.

For the experiment, we designed similar AI- and Human-interventions to encourage users to walk more than an *individual-specific* number of steps over seven days (i.e., one week). Each intervention reveals whether the step goal is generated by AI or health experts. We also designed Neutral-interventions that do not disclose the use of AI or health experts in generating the step goals. Table 2 presents the specific messages used for the three main interventions. We randomly assigned users into three groups (AI-, Human-, and Neutral-interventions) and sent a corresponding intervention message to each group.

Table 2. Main Intervention Messages

Group	Intervention Message
Neutral-interventions (Baseline)	Would you walk [*personalized step goal*] in the next seven days? Would you like to participate?
AI-interventions	AI recommends that you walk [*personalized step goal*] in the next seven days. Would you like to participate?
Human-interventions	Health expert recommends that you walk [*personalized step goal*] in the next seven days. Would you like to participate?

An intervention message contains an individual-specific step goal which is generated by a behavior prediction algorithm (see Aswani et al. 2019 and Zhou et al. 2018 for more details about the algorithm). Based on individuals' historical daily step data in the prior month, this AI algorithm leverages reinforcement learning to generate a challenging yet attainable step goal for the following seven-day period, which is predicted to maximize each user's expected step count in that period. We found that the step goals were generated in a reasonable manner; users with less (more) step count in the prior month were suggested to make a greater (smaller) percentage increase in their number of steps in the next week. It should be noted that step goals for the three intervention groups were generated by the same AI algorithm. This is to ensure that we focus on users' different *perceptions* (i.e., cognitive and affective trust) toward AI and health experts while controlling for the *actual performance* of AI and health experts.

We conducted our experiment during a one-week period in South Korea. For the experiment, we first randomly selected 3,000 unique users, who had used the focal app for more than one month, for each of the three intervention groups, resulting in a total of 9,000 unique users. Next, we calculated the step goal for each user using the AI algorithm. We then sent the interventions to the users through pop-up notifications of the app (e.g., see Figure 1). If a user tapped the "Yes" button in the pop-up notification, the app counted the user's number of steps for the next seven days. Our interventions did not involve any financial, social, or reputational incentives, and the focal app had never provided any interventions with personalized step goals for users before this experiment.

To test Hypothesis 1, users' acceptance of interventions is measured by outcome variable $Acceptance_i$, i.e., a binary variable indicating whether user *i* tapped the "Yes" button in the pop-up notification. To test Hypothesis 2, users' actual health behavior change after accepting interventions, or whether they made a sufficient effort and commitment to accomplish the recommended step goals, is measured by another outcome variable $Achievement_i$, i.e., a binary variable indicating whether user *i* walked more than her recommended step goal within seven days. By comparing the two outcome variables between the three intervention groups, we can identify the extent to which AI improves the focal firm's operational and business performance (i.e., how effectively AI replaces the extant human-based operations and promotes users' health behavior, respectively) in the context of preventive care.

After the experiment, we obtained additional information on users' gender, age, height, weight, previous walking activity, mobile app usage, and location. Information on gender, age, height, and weight was collected by in-app surveys before the experimental period, while that on previous walking activity, mobile app usage, and location was collected at the time of each user's receipt of an intervention. Accordingly, we could generate variables *Gender* (a binary variable where a value of one indicates female gender), *Age* (in years), *Height* (in centimeters), and *Weight* (in kilograms). In addition, *Previous Steps* indicates the number of steps walked in the past seven days, and *App Proficiency* represents how many times each user launched the focal app in the past seven days. Finally, *Location* indicates a series of dummy variables representing the 15 geographical areas in which each user received the intervention. Table 3 provides descriptive statistics.

Table 3. Descriptive Statistics for the Main Interventions

Group	Obs.	Accept. (%)	Achieve. (%)	Female (%)	Average (SD)				
					Age	Height	Weight	Previous Steps	App Proficiency
Neutral-interventions	3,000	11.33	5.86	54.05	51.05	164.67	64.52	49,359.24	3.373
				(0.49)	(13.90)	(10.15)	(17.15)	(31688.17)	(9.3527)
AI-interventions	3,000	18.70	10.30	53.53	50.97	164.51	65.59	49,758.91	3.585
				(0.49)	(13.74)	(10.95)	(17.27)	(31,989.82)	(9.321)
Human-interventions	3,000	22.36	13.23	54.03	51.04	164.08	64.54	50,177.93	3.507
				(0.49)	(13.89)	(10.94)	(17.61)	(37317.04)	(9.239)

To ensure the quality of the randomization procedure, we conducted an analysis of variance (ANOVA) and compared the means of the covariates (i.e., *Age*, *Gender*, *Height*, *Weight*, *Previous Steps*, *App Proficiency*, and *Location*) among the three groups. The results indicate that all the covariates other than *Previous Steps* are balanced and statistically similar across the three groups ($p > 0.1$). While *Previous Steps* is not balanced across the groups ($p < 0.1$), we obtained consistent results for the matched sample generated by using the propensity score matching (PSM) method (Rosenbaum and Rubin 1983).[5]

---

[5] The results for the matched sample are provided in Online Appendix (see Tables A2 – A4). We also adopted the difference-in-differences (DID) design to take account of the unobserved individual heterogeneity and found consistent results (see Section 5.2).

For the empirical analyses, we employed the Neutral-interventions group as the baseline group and model the probability or likelihood of whether each user accepted an assigned intervention and achieved an assigned goal [i.e., $Pr\ (Acceptance_i\ or\ Achievement_i = 1)$, respectively] as a logistic function of whether a user received AI- or Human-interventions as well as other control variables:

$$Pr\ (Acceptance_i\ or\ Achievement_i = 1 | AI_i, Human_i, X_i) = \frac{exp(U_i)}{1 + exp(U_i)}, \qquad (1)$$

$$U_i = \alpha + \beta_1 * AI_i + \beta_2 * Human_i + \tau * X_i + \varepsilon_i,$$

where *i* represents each user, and U_i denotes the latent utility of the intervention that user *i* received. AI_i and $Human_i$ are binary variables indicating which intervention user *i* received, i.e., AI- or Human-interventions. X_i is a vector of control variables regarding user characteristics, and ε_i comprises idiosyncratic error terms. To obtain less-biased estimates of users' different acceptance of AI- and Human-interventions, we controlled for the potential confounding effects of various user characteristics. Specifically, we incorporated *Age* and *Gender* to alleviate concern about the different effects of AI- and Human-interventions across people with different demographic characteristics. We also included *Height* and *Weight* to account for each user's physical characteristics, which could explain different perceptions of the preventive health interventions provided by AI and health experts. To account for each user's previous health behavior as well as her usual activity level, we included *Previous Steps* in our model. Given that higher proficiency in utilizing mobile health apps could lead to a higher propensity to accept a mobile intervention, we also incorporated *App Proficiency* in the model. In addition, we incorporated 14 *Location* dummies in the model to address the potential effect of geographical differences between users as well as contextual variations across different locations (e.g., weather).

Table 4 shows the results.[6] First, compared to the baseline Neutral-interventions group, the effects of AI-interventions are significantly positive regarding both outcome measures (i.e., $Acceptance_i$ and $Achievement_i$) ($p < 0.01$). This implies that the use and disclose of AI significantly improves the acceptance of preventive health interventions and individuals' health behavior, even after controlling for the potential effects of demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical characteristics (i.e., *Location*). This deviates from the previous findings on the negative impact of AI disclosure (e.g., Luo et al. 2019). Second, however, we find that the acceptance of AI-interventions is significantly lower than that of Human-interventions (i.e., for $Acceptance_i$, reject the null hypothesis: $\beta_{Human} - \beta_{AI} = 0$ at $p < 0.01$), thereby providing further support for Hypothesis 1. In addition, we identify the consistent result regarding individuals' actual health behavior

---

[6] The full results with the coefficient estimates for the control variables are provided in Online Appendix (see Table A22).

change. In other words, the effect of AI-interventions on *Achievement*$_i$ is significantly smaller than that of Human-interventions ($\beta_{Human} - \beta_{AI} = 0$ at $p < 0.01$), which supports Hypothesis 2. Specifically, compared to Neutral-interventions, AI-interventions (Human-interventions) increase the odds of *Acceptance* and *Achievement* by a factor of 1.80 (i.e., $e^{0.589}$) and 1.84 (2.27 and 2.45), respectively. The results align well with our theoretical arguments and survey results that lower affective trust in AI than health experts dominates its comparable cognitive trust and thereby results in lower acceptance of AI-interventions than Human-interventions. This corroborates the decisive role of affective trust in the acceptance of AI-interventions.

Table 4. Results of the Main Interventions

Variables	(1) Acceptance	(2) Achievement
AI	0.589***	0.609***
	(0.0743)	(0.0985)
Human	0.819***	0.896***
	(0.0724)	(0.0946)
Constant	–2.753***	–3.662***
	(0.569)	(0.856)
Controls	Y	Y
Observations	9,000	9,000
Log-likelihood	–4,089.02	–2,822.92

Note: The baseline group is the *Neutral-interventions* group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.2. How to Improve the Effectiveness of AI-Interventions

To test Hypotheses 3 and 4 regarding the roles of affective trust and thereby provide managerial implications for how to improve the effectiveness of AI-interventions, we extended our main survey and experiment with additional interventions. Specifically, we designed two additional interventions, each of which discloses the use of AI in tandem with health experts (i.e., AI-Human interventions) and provides a detailed explanation of how AI generated the interventions (i.e., AI-Transparency interventions). In other words, AI-Human and AI-Transparency interventions were designed to build different dimensions of affective trust in AI-interventions (i.e., benevolence and integrity, respectively) and thereby test Hypotheses 3 and 4, respectively. Thus, the AI-interventions group served as the baseline group in the following analyses. The specific messages used for the additional interventions are provided in Table 5.

First, using a survey, we investigated whether and to what extent AI-Human and AI-Transparency interventions improve affective trust in AI and consequent acceptance of AI-interventions. To this end, the survey questionnaires regarding AI-Human and AI-Transparency interventions were distributed together with the main interventions (i.e., AI- and Human-interventions). We collected 118 valid responses while discarding four low-quality samples.[7]

Table 5. Additional Intervention Messages

Group	Intervention Message
AI-interventions (Baseline)	AI recommends that you walk [*personalized step goal*][a] in the next seven days. Would you like to participate?
AI-Human interventions	By using AI, health expert recommends that you walk [*personalized step goal*] in the next seven days. Would you like to participate?
AI-Transparency interventions	AI recommends that you walk [*personalized step goal*] in the next seven days. AI has predicted how many steps you will walk in the next seven days based on your previous walking activity. Given that prediction, AI has selected a challenging yet attainable step goal that would maximize your physical activity. Would you like to participate?

Note: [a] For survey participants, the identical step goal (i.e., 49,776 steps) was provided.

Table 6 shows the survey results of the additional interventions. As can be seen in the table, compared to AI-interventions, AI-Human interventions improve benevolence by 1.350 (*t*-test, $p < 0.01$) while integrity is improved by 1.016 (*t*-test, $p < 0.01$). On the other hand, AI-Transparency interventions show 1.311 greater integrity compared with AI-interventions (*t*-test, $p < 0.01$), while their improvement in benevolence is 1.212 (*t*-test, $p < 0.01$). Thus, consistent with our theoretical arguments, the use of AI in tandem with human experts (providing transparent information on AI) is most helpful in improving benevolence (integrity) of AI, validating our experimental design. In addition, the results demonstrate that the AI-Human and AI-Transparency interventions groups show a significantly greater acceptance rate compared with the AI-interventions group (*t*-test, $p < 0.01$ and $p < 0.1$, respectively), lending support for Hypotheses 3 and 4, respectively. In other words, the inclusion of the additional features (i.e., human aspects, transparent information) into AI-interventions and the resulting increased affective trust in AI enhance the acceptance of AI-interventions.

Table 6. Additional Survey Results

Group	Obs.	Acceptance (%)	Benevolence	Integrity
			(Affective Trust)	
AI-interventions	49	34.70	2.484	2.479
AI-Human	62	59.68	3.834	3.495
AI-Transparency	56	48.21	3.696	3.790
Difference (AI-Human – AI)		24.98	1.350	1.016
p-value of *t*-test		$p < 0.01$	$p < 0.01$	$p < 0.01$
Difference (AI-Transparency – AI)		13.51	1.212	1.311
p-value of *t*-test		$p < 0.1$	$p < 0.01$	$p < 0.01$

To further identify whether the improved acceptance of AI-Human and AI-Transparency interventions holds true in practice and whether these additional interventions are also more effective for actual health

---

[7] The quality of the survey is described in detail in Online Appendix.

behavior change, we also extended the main experiment; each additional intervention was randomly sent to 3,000 unique users of the collaborating app without overlapping with the users who receive the main interventions. This results in a total of 6,000 additional unique users. Note that the additional interventions were sent on the same day in the same manner as the main interventions (i.e., AI-, Human-, Neutral-interventions). Table 7 shows descriptive statistics for the additional interventions. To ensure the quality of the randomization procedure, we conducted an ANOVA test and compared the means of the covariates among the three groups (i.e., AI-, AI-Human, and AI-Transparency interventions). The results indicate that all the observed attributes are statistically similar across the groups ($p > 0.1$).

Table 7. Descriptive Statistics for the Additional Interventions

Group	Obs.	Accept. (%)	Achieve. (%)	Female (%)	Average (SD)				
					Age	Height	Weight	Previous Steps	App Proficiency
AI-interventions	3,000	18.70	10.30	53.53 (0.49)	50.97 (13.74)	164.51 (10.95)	65.59 (17.27)	49,758.91 (31,989.82)	3.585 (9.321)
AI-Human	3,000	27.33	18.93	52.86 (0.49)	50.97 (14.08)	164.28 (10.48)	64.85 (16.61)	50,486.01 (31,411.84)	3.386 (9.126)
AI-Transparency	3,000	21.09	12.83	54.20 (0.49)	51.26 (13.63)	164.37 (10.58)	64.92 (16.75)	50,102.94 (33,861.77)	3.427 (8.876)

Table 8. Results of the Additional Interventions

Variables	(1) Acceptance	(2) Achievement
AI-Human	0.498***	0.718***
	(0.0625)	(0.0764)
AI-Transparency	0.212***	0.259***
	(0.0644)	(0.0816)
Constant	-1.514***	-1.678***
	(0.573)	(0.611)
Controls	Y	Y
Observations	9,000	9,000
Log-likelihood	-3582.874	-4779.8793

Note: The baseline group is the AI-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We replicated the main analysis (i.e., Equation 1) for the users with AI-Human, AI-Transparency, and AI-interventions, utilizing the AI-interventions group as the baseline group. Table 8 shows the results.[8] Given that the AI-interventions group was used as the baseline group, a significant and positive coefficient of each of the additional interventions (e.g., AI-Human, AI-Transparency) implies its greater effectiveness than AI-interventions. As can be seen in the first column of Table 8, users' acceptance of AI-Human interventions is significantly greater than that of AI-interventions ($p < 0.01$). AI-Transparency interventions also exhibit a significantly greater acceptance compared with AI-interventions ($p < 0.01$). The results are

---

[8] The full results with the coefficient estimates for the control variables are provided in Online Appendix (see Table A23).

also consistent with users' actual health behavior change (i.e., *Achievement*, see the second column of Table 8). Specifically, compared to AI-interventions, AI-Human (AI-Transparency) interventions increase the odds of *Acceptance* and *Achievement* by a factor of 1.65 (i.e., $e^{0.498}$) and 2.05 (1.24 and 1.30), respectively. The results collectively support Hypotheses 3 and 4.

These findings of the additional interventions demonstrate that improving human aspects or transparency of AI-interventions significantly enhances deficient aspects of affective trust (e.g., benevolence, integrity) in AI and in turn increases the effectiveness of AI-interventions. Thus, the results of the additional interventions further corroborate that affective trust plays a decisive role in the acceptance of AI-interventions.

5. Additional Analyses

5.1. Causality between Affective Trust and Acceptance of AI-Interventions

To provide more direct evidence for the causal relationship between affective trust in AI and acceptance of AI-interventions, we conducted an additional experiment.[9] To this end, we collaborated with Macromill Embrain, a market research company that specializes in consumer research and survey with 6.4 million panels worldwide. In the additional experiment, participants were randomly assigned into a treatment or control group. For the control group, we provided a simple description of the focal app and its AI-based preventive health intervention. On the other hand, the treatment group was offered additional descriptions designed to improve affective trust in AI. Specifically, the two sets of additional descriptions were provided to improve benevolence and integrity of AI by emphasizing its genuine care and warmth as well as transparency, respectively. Table 9 shows the specific descriptions given to the control and treatment groups. After providing the different descriptions to the control and treatment groups, we measured their affective trust in AI. We then showed the identical AI-interventions to both groups (see the left side of Figure 1) and measured their acceptance.

For the experiment, we recruited 85 and 86 participants for the control and treatment groups, respectively. Among 171 participants, 49.7% are female, and their average age is 39.4. We conducted *t*-tests and found that the treatment and control groups are statistically similar in age ($p > 0.1$) and gender ($p > 0.1$). To identify whether the additional information given only to the treatment group actually improves their affective trust in AI, we compared the mean values of each dimension of affective trust (i.e., benevolence and integrity) between the treatment and control groups. The result shows that the treatment group exhibits higher mean benevolence and integrity than the control group (*t*-test: $p < 0.01$). Such higher mean affective trust validates our experimental design. We then examined whether this increased affective

---

[9] We thank the anonymous reviewer for this valuable suggestion.

trust in AI actually results in greater acceptance of AI-interventions. To this end, we compared the acceptance rates of AI-interventions between the treatment and control groups. The result demonstrates that the treatment group shows a significantly higher acceptance rate than the control group (t-test: $p < 0.1$). The results further confirm that an increase in affective trust in AI causes greater acceptance of AI-interventions.

Table 9. Descriptions Given to Different Experimental Groups

Group	Description
Control	Suppose that you are using a mobile healthcare application, wherein its AI (Artificial Intelligence) provides you with a personalized health recommendation about the number of walking steps for the next week based on your number of steps in the last month.
Treatment	Suppose that you are using a mobile healthcare application, wherein its AI (Artificial Intelligence) provides you with a personalized health recommendation about the number of walking steps for the next week based on your number of steps in the last month. In this context, our AI, named AIden, is taking good care of you and try very hard to provide just the right recommendation to you. AIden is always ready to answer any questions you might have about suggested recommendations and their impact on your health. In addition, AIden is genuine and believable as it provides detailed and easy explanation about how recommended goals are generated. For example, AIden has predicted how many steps you will walk in the next seven days based on your previous walking activity. Given that prediction, AIden has selected a challenging yet attainable step goal that would maximize your physical activity.

5.2. Actual Health Behavior Change

To investigate individuals' health behavior change caused by interventions, we focused on *whether* users achieved recommended step goals or not (i.e., *Achievement*) in the main analyses. However, another important outcome regarding AI for preventive care operations is *to what extent* it improves users' actual health behavior. Though we have shown the significant differences in the achievement rate of the different interventions, the practical implications would diminish if the difference disappeared in terms of the *amount* of health behavior change. To investigate the amount of users' actual health behavior change driven by the different interventions, we introduce an additional continuous outcome variable, i.e., $Average\ Steps_{i,t}$, which indicates user i's average daily number of steps walked in week t. Given the strict data protection policy of the firm, we could obtain each user's $Average\ Steps_{i,t}$ only in the week before the treatments (i.e., t = -1) as well as the first, second, third, and sixth weeks after the treatments (i.e., t = 1, 2, 3, and 6). We applied the log-transformation to $Average\ Steps_{i,t}$ to improve its normality. We constituted the panel data consisting of users' weekly number of steps walked before and after the interventions. Accordingly, the panel data set includes 45,000 observations: 3 groups (i.e., Neutral-, AI-, and Human-interventions) × 3,000 users × 5 weeks (i.e., 1 pre-treatment and 4 post-treatment weeks). We then applied the difference-in-differences (DID) method with individual- and week-fixed effects to the Neutral-, AI-, and Human-interventions groups, while considering the Neutral-interventions group as the control and the AI- and Human-interventions groups as the two different treatments:

$$\begin{aligned} Average\ Steps_{i,t} = \beta_0 + \beta_1 * AfterIntervention_t \\ +\beta_2 * AI_i * AfterIntervention_t \\ +\beta_3 * Human_i * AfterIntervention_t + \alpha_i + \mu_t + \varepsilon_{i,t}, \end{aligned} \quad (2)$$

where *i* indicates each user, and *t* denotes each week (i.e., -1, 1, 2, 3, 6). $AfterIntervention_t$ is a dummy variable indicating whether t is before (i.e., t = -1) or after the interventions (i.e., t = 1, 2, 3, or 6). The individual-fixed effect, α_i, allows us to account for the unobserved time-invariant individual heterogeneity. μ_t is the week-fixed effect which controls weekly variation in the average daily number of steps. The parameters of interest are the coefficients of the interaction terms, i.e., β_2 and β_3. Specifically, β_2 (β_3) indicates the difference between the effects of Neutral- and AI-interventions (Human-interventions) on average daily walking steps.

The results are provided in the first panel of Table 10. The significant positive values of $\hat{\beta}_2$ and $\hat{\beta}_3$ ($p < 0.01$) indicate that AI- and Human-interventions lead to more health behavior change than Neutral-interventions, respectively. In addition, a significantly smaller estimated value of β_2 than β_3 ($p < 0.01$) shows that AI-interventions induce less health behavior change than Human-interventions, supporting Hypothesis 2 further. Specifically, compared to Neutral-interventions, AI- and Human-interventions increase *Average Steps* by 12.5% and 32.8% (i.e., $e^{0.118} - 1$ and $e^{0.284} - 1$), respectively.

Table 10. Results of DID Analyses

(1) Main Interventions		**(2) Additional Interventions**	
Variables	**Average Steps**	**Variables**	**Average Steps**
AfterIntervention (β_1)	0.108*** (0.0257)	AfterIntervention (β_1)	0.729*** (0.0361)
AI * AfterIntervention (β_2)	0.118*** (0.0309)	AI-Human *AfterIntervention (β_2)	0.326*** (0.0579)
Human * AfterIntervention (β_3)	0.284*** (0.0202)	AI-Transparancy * AfterIntervention (β_3)	0.150*** (0.0513)
Constant	8.597*** (0.00926)	Constant	8.599*** (0.0182)
Individual and Week Fixed Effects	Y	Individual and Week Fixed Effects	Y
Observations	45,000	Observations	45,000
R^2	0.0083	R^2	0.0021

Note: Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We also replicated the same analyses for the additional interventions. Specifically, we applied the same DID specification with individual- and week-fixed effects to the AI-, AI-Human, and AI-Transparency interventions groups. Therefore, users with AI-interventions are considered the control group, while those with AI-Human and AI-Transparency interventions represent the two different treatment groups. The second panel of Table 10 shows the results, which are consistent with the main results. AI-Human and AI-

Transparency interventions lead to more health behavior change than AI-interventions ($p < 0.01$), further supporting Hypotheses 3 and 4, respectively. Specifically, compared to AI-interventions, AI-Human and AI-Transparency interventions increase *Average Steps* by 38.5% and 16.2%, respectively. Thus, the results of DID analyses collectively show that our findings remain valid for the amount of health behavior change and are robust against the unobserved time-invariant individual heterogeneity and time-specific peculiarities.

5.3. Comparison with Human-interventions

To further examine whether properly designed AI-based interventions can be even more effective than Human-interventions, we compare AI-Transparency and AI-Human interventions, which showed greater acceptance than AI-interventions, to another control group, i.e., Human-interventions. Specifically, we replicated the main analyses for the users with AI-Human, AI-Transparency, and Human-interventions, while utilizing the Human-interventions group as the baseline group (see Table 11 for the results). We observe that the difference in the effectiveness between AI-Transparency and Human-interventions is statistically insignificant ($p > 0.1$). This result stresses that improving transparency of AI-interventions helps compensate for the affective deficiency in AI, making the effectiveness of AI-Transparency interventions comparable to that of human-based interventions. Moreover, the use of AI in tandem with human experts (i.e., AI-Human interventions) makes AI-interventions even more effective than Human-interventions ($p < 0.01$). This result implies that supplementing Human-intervention with AI can be an attractive strategy to derive the greatest effectiveness. This additional comparison provides further implications for how to make AI-based interventions comparable to or even more effective than Human-interventions.

Table 11. Results of the Additional Interventions with Human-Interventions as the Alternative Baseline Group

Variables	(1) Acceptance	(2) Achievement
AI-Transparency	0.0165	0.0337
	(0.0622)	(0.0772)
AI-Human	0.283***	0.361***
	(0.0604)	(0.0722)
Constant	–1.507***	-2.364***
	(0.579)	(0.647)
Controls	Y	Y
Observations	9,000	9,000
Log-likelihood	–4,928.15	-3758.0585

Note: The baseline group is the Human-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.4. Robustness Checks

We conducted a series of additional analyses to address potential concerns regarding the results. A summary of potential threats to our results and corresponding additional analyses is provided in Table 12. The main findings are robust against all the different factors that we considered. The details of the robustness checks are provided in Online Appendix.

Table 12. Summary of Robustness Checks

Potential Concerns	Additional Analyses	Appendix Location
Previous Steps of the Human-interventions group was slightly higher than that of the AI- and Neutral-interventions groups.	Replicated the main analysis by adopting propensity score matching (PSM) method	Tables A2 – A4
Achievement could be captured only if *Acceptance* = 1.	Replicated the analysis of *Achievement* only for users whose *Acceptance* = 1	Table A5
	Adopted Heckman (1976)'s selection model	Tables A6 – A7
The difference between the previous number of steps and the suggested step goal could affect the effectiveness of different interventions.	Replicated the main analyses with an additional control variable which indicates the gap between the previous number of steps and the suggested step goal	Tables A8 – A9
Given that the term *experts* could be related to authority or social obligation, the results could be different if a softer message was used in Human-interventions.	Conducted an additional experiment to identify whether the use of a softer message results in lower acceptance of Human-interventions	Table A10
The use of an additional question mark only in Neutral-interventions would affect individuals' responsiveness to Neutral-interventions.	Replicated the main analysis while using an alternative control group of users who did not receive any interventions	Table A11
The results could be biased by users who did not read the intervention messages.	Replicated the main analyses for users who can be considered to read the intervention messages	Tables A12 – A14
The results could not be generalizable toward people in different age groups.	Replicated the main analyses for users in three different age groups (i.e., Under 30, 30 – 59, and Over 60)	Tables A15 – A18
The results could be biased by people who already walked enough before the experiments.	Replicated the main analyses for users whose *Previous Steps* are less than a certain level	Tables A19 – A20
The results could be explained by greater "uniqueness neglect" of AI than humans, which is identified as a critical factor of AI acceptance in a diagnostic care context.	Conducted a survey to measure individuals' perceived "uniqueness neglect" toward AI- and Human-interventions	Table A21

6. Discussion

6.1. Theoretical Contributions

The specific theoretical contributions of this study are twofold. First, this study contributes to the literature on the role of new-age technologies in behavioral operations management (e.g., Donohue et al. 2020) and consumer marketing (e.g., Kumar 2021). The development of healthcare AI has concentrated mainly on the context of diagnostic care, where healthcare providers directly drive its adoption. Accordingly, considerable research has been devoted to how healthcare providers perceive and adopt AI-based services in the diagnostic care context (e.g., AI-based disease diagnosis and treatment) (Krittanawong 2018, Palanica et al. 2019). In preventive care, however, laypeople are empowered to interact directly with AI-based preventive health services and make acceptance decisions without being much influenced by healthcare providers (Kwon et al. 2022, Yu et al. 2018). Thus, lay consumers' high receptivity to AI-based preventive health services is critical for preventive care providers to achieve their enhanced operational performance, business outcomes, and public health. Nevertheless, while a number of AI-based preventive health services are readily available in the market and draw considerable attention from preventive care providers who seek to improve their operational performance and consumer behavior, little has been known about how the general public perceives and accepts AI-based preventive health services as well as how they change their actual health behavior accordingly. Drawing on a randomized field experiment complemented by a survey, this study responds to this call.

While we examined the acceptance of AI-based services in a specific context, i.e., preventive care, our results can be extended to other contexts where affective trust plays a significant role, such as travel, education, legal, and insurance services (Alford and Sherrell 1996, Patti and Chen 2009). Similar to preventive care, these industries are often considered a credence-based field, where consumers find it difficult to evaluate the exact performance of a service after consumption due to the lack of objective evidence and necessary knowledge to evaluate the performance as well as the long-time gap until outcomes are realized (Alford and Sherrell 1996, Darby and Karni 1973). As a result, in a credence-based field, consumers rely heavily on providers' affective signals (Darby and Karni 1973). Thus, as we theorized, such a significant role of affective trust in a credence-based field, together with the low affective trust in AI in general, would result in the low acceptance of AI-based services compared with their human counterparts.

Second, this study contributes to the technology acceptance and trust literature by directly comparing the effectiveness of AI- and Human-interventions as well as assessing the roles of cognitive and affective trust in the acceptance of AI-interventions. As AI is expected to replace the tasks that have been typically conducted by humans, it is increasingly critical to understand what factors drive users to accept AI-based services *over* human-based ones, or vice versa. However, while trust has been suggested as a key challenge regarding users' acceptance of AI (Kumar 2021), prior technology acceptance studies did not provide comprehensive insights into this issue. This is because these studies exclusively considered technology applications and examined the role of different features in their acceptance, without much consideration of the comparable tasks conducted by humans. Accordingly, the previous studies emphasized the rational aspect of trust, which cannot explain the current lower overall trust in AI-interventions than Human-interventions even though AI-interventions often show comparable performance with or even outperform Human-interventions. Thus, this makes it difficult to theoretically explain or predict the relative acceptance of AI- and Human-interventions. By broadening the scope of trust to include the affective aspect, which could be divorced from objective performance, this study revises the extant theoretical perspective to explain the lower acceptance of AI-interventions, particularly in a context where affective trust plays a decisive role (e.g., credence-based fields). Based on this theoretical perspective, this study also provides theory-driven, practically validated strategies to enhance the extent to which AI-interventions improve operational and business performance in the preventive care context. Specifically, we identified that including human-like features, providing an explicit explanation of how AI generates intervention, and highlighting its genuine care and warmth, results in greater acceptance of AI-interventions by improving affective trust in AI.

6.2. Managerial Implications

This study also provides valuable practical implications for how firms should design and exploit their AI applications regarding preventive care operations. First, the higher effectiveness of AI-interventions compared with Neutral-interventions underscores the positive effect of highlighting the use of AI on encouraging more intervention acceptance and health behavior change. While various AI algorithms are being used to provide personalized services and interventions for preventive care operations, the use of AI is rarely explicitly revealed because such information is often considered irrelevant. Previous findings on the negative effect of AI disclosure (e.g., Luo et al. 2019) also contribute to this trend. However, our results suggest that such practice misses the opportunity to further improve the operational performance and ultimately enhance public health. Thus, we recommend that the use of AI in preventive health interventions should be clearly disclosed. Nevertheless, the replacement of existing Human-interventions with AI-interventions should be approached carefully. This is because our results indicate that users receiving AI-interventions show less acceptance and health behavior change compared with those receiving Human-interventions even though both interventions were generated by the same AI algorithm.

This study also calls for managerial efforts to reduce users' resistance to AI through active customer communication and to make AI applications more trustworthy. Specifically, the results collectively illustrate the role of trust in the acceptance of AI-interventions. Thus, firms should be cautious to avoid the common pitfall of believing that improving the objective performance of their technology is sufficient to convince consumers to use the technology and thus neglecting efforts to build consumer trust through effective communications and personalized attention (Slater and Mohr 2006). According to our results, budgets should be allocated appropriately to strike a balance between the technological advancements of and users' receptivity to AI.

Moreover, our theoretical arguments and survey results consistently demonstrate that such low trust in AI is rooted in its low affective trust rather than cognitive trust. Despite the recent debate on the importance of affective aspects (e.g., warmth, kindness, and humanism) in users' trust and positive responses to AI, we do not see many practitioners tapping affective aspects directly in their AI-based services. Having the causal understanding that the affective aspects of AI drive the effectiveness of AI-interventions, preventive care providers should consider facilitating affective trust in their AI-based preventive health interventions in general, physical activity promotion in particular. Specifically, the results demonstrate that designing AI in a way to highlight its genuine care and warmth would improve its acceptance by enhancing affective trust. More generally, given that affective trust consists of benevolence and integrity, providers could improve the effectiveness of AI-based preventive health interventions by cultivating their image that they carefully consider consumers, have good intentions toward consumers, put consumers' interests before their own (i.e., benevolent), and are reliable and honest (i.e., integrity).

Our results also provide more specific implications for how AI-based preventive health interventions can be designed in a more affectively trustworthy and effective way. First, the finding that AI-Human interventions are more effective than AI- and Human-interventions strongly recommends exploiting AI-interventions together with human experts. For example, an ideal scenario would be letting existing health experts use AI to generate more effective preventive health interventions with greater scalability, while explicitly emphasizing the involvement and opinions of human experts as well as the use of AI. Increasing number of preventive care service providers, including popular weight loss app Noom Coach, have implemented such an approach. This further substantiates the perspective that AI would have the most significant effect when it augments humans rather than replacing them (Wilson and Daugherty 2018).

The results also stress the important role of transparency in building trust in AI and improving its consequent acceptance, a recommendation that is consistent with the growing emphasis on the notion of explainable AI (XAI) (Anjomshoae et al. 2019). This is also well aligned with the literature on operational transparency (e.g., Buell et al. 2021, Lee et al. 2021). Given that the acceptance of AI-interventions in our study was improved when information about the underlying mechanisms behind the recommended personalized goals was provided, AI-based preventive health interventions should be offered with an explicit explanation of how AI generates their services or interventions (e.g., AI-Transparency interventions). Providers offering human-based preventive health interventions (e.g., Human-interventions) can also consider adopting transparent AI-based interventions (e.g., AI-Transparency interventions) given that Human- and AI-Transparency interventions have statistically similar effectiveness. Moreover, governments and regulatory agencies need to establish a legal framework for transparent AI, especially in the preventive care sector, in order to improve the effects of AI-based preventive health services on public health as well as consumers' right to receive an explanation for an algorithm-based decision. For example, the European Union's new General Data Protection Regulation (GDPR) requires businesses to provide information on the logic of AI-based decision-making processes (Wallace and Castro 2018).

6.3. Limitations and Future Research

Several limitations of the study are noteworthy and pave the way for future research. First, the sample in our experiments is a set of users who voluntarily downloaded and installed the focal healthcare app. Thus, while we believe our sample represents users with the real motivations that drive health behavior (Baek and Shore 2020), at the same time, it is also possible that our sample represents people with a specific characteristic. Though we endeavored to address this concern by controlling for their proficiency with mobile healthcare apps, we acknowledge that this deficiency may not have been fully resolved. Second, unobserved time-variant individual heterogeneity might threaten our identification strategy. While we conducted various robustness checks to address individual heterogeneity, the cross-sectional nature of this study does not allow us to completely rule out the potential confounding effect of unobserved individual heterogeneity. Another caveat is related to the recommended step goals. While our AI algorithm considered users' previous health behavior (i.e., step records) to generate step goals, more sophisticated health recommendations would take account of their diverse health characteristics. We call for future studies that would address these concerns and extend the validity of our findings. In addition, future research could conduct additional experiments with the full 2 x 2 x 2 factorial design to investigate how additional features designed to improve *either* affective or cognitive trust affect individuals' acceptance of AI- and Human-interventions differently, which can provide a more comprehensive understanding of trust in and acceptance of AI for preventive care. Lastly, this study focused on a specific preventive care context where interventions are designed to promote physical activity. While we provided the theoretical arguments and a series of anecdotal evidence of low affective trust in AI in general preventive care contexts, our findings might not be generalizable into a particular preventive care context where individuals have higher affective trust in AI than human counterparts. Future studies can explore such contexts and investigate the boundary conditions for our findings.

7. Conclusions

Our results from a randomized field experiment with 15,000 unique users of a mobile healthcare app show that disclosing the use of AI in preventive health interventions (i.e., AI-interventions) induces more acceptance and health behavior change than providing the interventions without disclosure (i.e., Neutral-interventions). However, the effectiveness of AI-interventions is lower than that of Human-interventions consistently across the different analyses. The survey results show that the lower acceptance of AI-interventions is attributable to individuals' lower affective trust in AI than health experts. In addition, the effectiveness of AI-interventions is improved when the affective trust features (i.e., human collaboration, transparency, genuine care, and warmth) are successfully added. These results collectively substantiate the perspective that affective trust, rather than typical cognitive trust, plays a decisive role in the acceptance of AI-interventions.

Given the escalating demand for and constrained supply of manpower for preventive care operations, AI has the potential to address this gap and improve the operational and business performance of preventive care providers by offering more effective preventive health interventions in place of health experts at a lower cost. Given that the general public is the ultimate consumer of AI for preventive care operations, understanding how they perceive and accept it is crucial for its effective utilization. However, users' reluctance to trust AI for preventive care operations despite its rapid technological development warns that the unconditional replacement of health experts with AI would undermine firms' performance. This becomes a more serious issue as firms are increasingly required by consumers and regulations to reveal whether and how they use AI in their services and products.[10] Nevertheless, due to the clear distinction between AI and conventional technologies, little is known about the unique aspects of individuals' acceptance of AI for preventive care operations. Motivated by this gap, this study aims to identify how and why people accept AI- and Human-interventions differently, thereby providing a solid theoretical framework to explain the unique aspects of the acceptance of AI for preventive care operations, which deviate from the acceptance of conventional technologies, as well as managerial implications for more effective use of AI for preventive care operations. We hope this research motivates additional inquiries into how we could extend and revise our accumulated knowledge in order to manage threats and opportunities for this new-age technology.

Acknowledgement

The authors thank the special issue editors and anonymous reviewers for their valuable and constructive suggestions that significantly improved this research. We also appreciate the helpful comments and suggestions from Sanghee Lim about the initial draft of this work. We would like to express our sincere gratitude to Swallaby Inc. and Haekwon Peter Chung for their support in the experiments and data analysis. This research is supported by the Nanyang Technological University, Singapore, under its Start-Up-Grant

---

[10] Given that health-related decisions are often associated with substantial risk (Epstein and Peters 2009), consumers become more likely to establish a clear-cut line of responsibility for the use of AI in healthcare services or products. In response to such consumer needs, regulations and policies increasingly compel businesses to reveal and explain the use of AI in their services or products. For example, the European Union's new General Data Protection Regulation (GDPR) requires businesses to provide information on the logic of AI-based decision-making processes (Wallace and Castro 2018). Such regulatory pressure is more pronounced in industries where more sensitive personal data is utilized, such as healthcare.

(SUG).

References

Alford, B. L., D. L. Sherrell. 1996. The role of affect in consumer satisfaction judgments of credence- based services. *J. Bus. Res.* **37**(1): 71–84.

Anjomshoae, S., A. Najjar, D. Calvaresi, K. Främling. 2019. Explainable agents and robots: Results from a systematic literature review. Proceeding of the 18th International Conference on Autonomous Agents and Multiagent Systems, May 13–17, Montreal, Canada.

Aral, S., D. Walker. 2011. Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Manag. Sci.* 57(9): 1623-1639.

Aswani, A., P. Kaminsky, Y. Mintz, E. Flowers, Y. Fukuoka. 2019. Behavioral modeling in weight loss interventions. *Eur. J. Oper. Res.* **272**(3): 1058–1072.

Baek, J., J. Shore. 2020. Forum size and content contribution per person: A field experiment. *Manag. Sci.* **66**(12): 5906–5924.

Beaton, T. 2017. How preventive healthcare services reduce spending for payers. Available at: https://healthpayerintelligence.com/news/how-preventive-healthcare-services-reduce-spending-for-payers (accessed date: October 16, 2021)

Bendoly, E. 2013. Real-time feedback and booking behavior in the hospitality industry: Moderating the balance between imperfect judgment and imperfect prescription. *J. Oper. Manag.* **31**(1–2): 62–71.

Borsky, A., C. Zhan, T. Miller, Q. Ngo-Metzger, A. S. Bierman, D. Meyers. 2018. Few Americans receive all high-priority, appropriate clinical preventive services. *Health Aff.* **37**(6): 925–928.

Brynjolfsson, E., A. Mcafee. 2017. Artificial intelligence, for real. *Harv. Bus. Rev.*

Buell, R. W., E. Porter, M. I. Norton. 2021. Surfacing the submerged state: operational transparency increases trust in and engagement with government. *Manuf. Serv. Oper. Manag.* **23**(4): 781–802.

Camerer, C. F., G. Nave, A. Smith. 2019. Dynamic unstructured bargaining with private information: theory, experiment, and outcome prediction via machine learning. *Manag. Sci.* **65**(4): 1867–1890.

Castelo, N., M. W. Bos, D. R. Lehmann. 2019. Task-dependent algorithm aversion. *J. Mark. Res.* **56**(5): 809–825.

Chiles, T. H., J. F. McMackin. 1996. Integrating variable risk preferences, trust, and transaction cost economics. *Acad. Manage. Rev.* **21**(1): 73–99.

Choudhary, V., M. Shunko, S. Netessine. 2021. Does immediate feedback make you not try as hard? A study on automotive telematics. *Manuf. Serv. Oper. Manag.* **23**(4): 835–853.

Collier, M., R. Fu, L. Yin, P. Christiansen. 2017. Artificial intelligence: healthcare's new nervous system. Available at: https://www.ehidc.org/resources/artificial-intelligence-healthcare%E2%80%99s-new-nervous-system (accessed date: October 16, 2021)

Cui, R., M. Li, S. Zhang. 2022. AI and Procurement. *Manuf. Serv. Oper. Manag.* **24**(2): 691–706.

Cummings, L. L., P. Bromiley. 1996. The organizational trust inventory (OTI). *Trust Organ. Front. Theory Res.* **302**(330): 39–52.

Darby, M. R., E. Karni. 1973. Free competition and the optimal amount of fraud. *J. Law Econ.* **16**(1): 67–88.

Davis, F. D. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **13**(3): 319–340.

Dawes, R. M. 1979. The robust beauty of improper linear models in decision making. *Am. Psychol.* 571–582.

Doney, P. M., J. P. Cannon. 1997. An examination of the nature of trust in buyer–seller relationships. *J. Mark.* **61**(2): 35–51.

Donohue, K., Ö. Özer, Y. Zheng. 2020. Behavioral operations: Past, present, and future. *Manuf. Serv. Oper. Manag.* **22**(1): 191–202.

Douglas, F., N. Torrance, E. Van Teijlingen, S. Meloni, A. Kerr. 2006. Primary care staff's views and experiences related to routinely advising patients about physical activity. A questionnaire survey. *BMC*

Public Health **6**(1): 138.

Ellis, S. C., S. Rao, D. Raju, T. J. Goldsby. 2018. RFID tag performance: linking the laboratory to the field through unsupervised learning. *Prod. Oper. Manag.* **27**(10): 1834–1848.

Epstein, R. M., E. Peters. 2009. Beyond information: exploring patients' preferences. *Jama* **302**(2): 195–197.

Fan, W., J. Liu, S. Zhu, P. M. Pardalos. 2020. Investigating the impacting factors for the healthcare professionals to adopt artificial intelligence-based medical diagnosis support system (AIMDSS). *Ann. Oper. Res.* **294**(1): 567–592.

Gartner. 2018. Gartner says global artificial intelligence business value to reach $1.2 trillion in 2018.Avaialbe at: https://www.gartner.com/en/newsroom/press-releases/2018-04-25 (accessed date: October 16, 2021).

Gatling, A., C. Shum, L. Book, B. Bai. 2017. The influence of hospitality leaders' relational transparency on followers' trust and deviance behaviors: mediating role of behavioral integrity. *Int. J. Hosp. Manag.* **62**: 11–20.

Gefen, D. 2002. Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM SIGMIS Database DATABASE Adv. Inf. Syst.* **33**(3): 38–53.

Gefen, D., E. Karahanna, D. W. Straub. 2003. Trust and TAM in online shopping: An integrated model. *MIS Q.* **27**(1): 51–90.

Gefen, D., D. W. Straub. 2004. Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. *Omega* **32**(6): 407–424.

Grove, W. M., D. H. Zald, B. S. Lebow, B. E. Snitz, C. Nelson. 2000. Clinical versus mechanical prediction: a meta-analysis. *Psychol. Assess.* **12**(1): 19.

Guha, S., S. Kumar. 2018. Emergence of big data research in operations management, information systems, and healthcare: past contributions and future roadmap. *Prod. Oper. Manag.* **27**(9): 1724–1735.

Gursoy, D., O. H. Chi, L. Lu, R. Nunkoo. 2019. Consumers acceptance of artificially intelligent (AI) device use in service delivery. *Int. J. Inf. Manag.* **49**: 157–169.

Ha, B., Y. Park, S. Cho. 2011. Suppliers' affective trust and trust in competency in buyers: Its effect on collaboration and logistics efficiency. *Int. J. Oper. Prod. Manag.* **31**(1): 56–77.

Halpern, J. 2003. What is clinical empathy? *J. Gen. Intern. Med.* **18**(8): 670–674.

Haslam, N., Y. Kashima, S. Loughnan, J. Shi, C. Suitner. 2008. Subhuman, inhuman, and superhuman: Contrasting humans with nonhumans in three cultures. *Soc. Cogn.* **26**(2): 248–258.

Hoejmose, S., S. Brammer, A. Millington. 2012. "Green" supply chain management: The role of trust and top management in B2B and B2C markets. *Ind. Mark. Manag.* **41**: 609–620.

Hopp, W. J., J. Li, G. Wang. 2018. Big data and the precision medicine revolution. *Prod. Oper. Manag.* **27**(9): 1647–1664.

Hurling, R., M. Catt, M. D. Boni, B. Fairley, T. Hurst, P. Murray, A. Richardson, J. Sodhi. 2007. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. *J. Med. Internet Res.* **9**(2): e633.

Ibanez, M. R., B. R. Staats. 2018. Behavioral empirics and field experiments. *Handb. Behav. Oper.* 121–147.

Johnson, D., K. Grayson. 2005. Cognitive and affective trust in service relationships. *J. Bus. Res.* **58**(4): 500–507.

Johnson, M., A. Albizri, S. Simsek. 2020. Artificial intelligence in healthcare operations to enhance treatment outcomes: a framework to predict lung cancer prognosis. *Ann. Oper. Res.* **308:** 275–305.

Kim, S.-H., J. Tong, C. Peden. 2020. Admission control biases in hospital unit capacity management: How occupancy information hurdles and decision noise impact utilization. *Manag. Sci.* **66**(11): 5151–5170.

Komiak, S. Y., I. Benbasat. 2006. The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Q.* 941–960.

Kong, F., J. Tan. 2012. DietCam: Automatic dietary assessment with mobile camera phones. *Pervasive Mob. Comput.* **8**(1): 147–163.

Kreuzfeld, S., M. Preuss, M. Weippert, R. Stoll. 2013. Health effects and acceptance of a physical activity program for older long-term unemployed workers. *Int. Arch. Occup. Environ. Health* **86**(1): 99–105.

Krittanawong, C. 2018. The rise of artificial intelligence and the uncertain future for physicians. *Eur. J. Intern. Med.* **48**: e13–e14.

Kumar, V. 2021. *Intelligent Marketing: Employing New-Age Technologies.* SAGE Publishing, New Delhi, India.

Kumar, V., S. Kumar, R. P. Leone. 2021. Transformative marketing and operations management. *Prod. Oper. Manag.* **30**(1): 295–297.

Kwon, H. E., Dewan, S., Oh, W., Kim, T. 2022. Self-Regulation and External Influence: The Relative Efficacy of Mobile Apps and Offline Channels for Personal Weight Management. *Inf. Syst. Res.* Forthcoming.

LaRosa, E., D. Danks. 2018. Impacts on trust of healthcare ai. Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, February 2-3, New Orleans, USA.

Lee, J., H. S. (Huck) Lee, H. Shin, V. Krishnan. 2021. Alleviating drug shortages: the role of mandated reporting induced operational transparency. *Manag. Sci.* **67**(4): 2326–2339.

Lewis, J. D., A. Weigert. 1985. Trust as a social reality. *Soc. Forces* **63**(4): 967–985.

Logg, J. M., J. A. Minson, D. A. Moore. 2019. Algorithm appreciation: People prefer algorithmic to human judgment. *Organ. Behav. Hum. Decis. Process.* **151**: 90–103.

Longoni, C., A. Bonezzi, C. K. Morewedge. 2019. Resistance to medical artificial intelligence. *J. Consum. Res.* **46**(4): 629–650.

Luo, X., S. Tong, Z. Fang, Z. Qu. 2019. Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Mkt. Sci*. **38**(6): 937-947.

Nelson, C. A., L. M. Pérez-Chada, A. Creadore, S. J. Li, K. Lo, P. Manjaly, A. B. Pournamdari, E. Tkachenko, J. S. Barbieri, J. M. Ko. 2020. Patient perspectives on the use of artificial intelligence for skin cancer screening: a qualitative study. *JAMA Dermatol*. **156**(5): 501–512.

Norman, P., Conner, M., and Bell, R. 1999. The theory of planned behavior and smoking cessation. *Health Psychol.* **18**(1):89.

Martelaro, N., V. C. Nneji, W. Ju, P. Hinds. 2016. Tell me more designing HRI to encourage more trust, disclosure, and companionship. Proceeding of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction, Daegu, South Korea.

Mayer, R. C., J. H. Davis, F. D. Schoorman. 1995. An integrative model of organizational trust. *Acad. Manage. Rev.* **20**(3): 709–734.

McAllister, D. J. 1995. Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. *Acad. Manage. J.* **38**(1): 24–59.

McCoy, P., S. Leggett, A. Bhuiyan, D. Brown, P. Frye, B. Williams. 2017. Text messaging: an intervention to increase physical activity among African American participants in a faith-based, competitive weight loss program. *Int. J. Environ. Res. Public. Health* **14**(4): 326.

Morgan, R. M., S. D. Hunt. 1994. The commitment-trust theory of relationship marketing. *J. Mark.* **58**(3): 20–38.

Nguyen, Q., T. H. Kim. 2019. Promoting adoption of management practices from the outside: Insights from a randomized field experiment. *J. Oper. Manag.* **65**(1): 48–61.

Nicolaou, A. I., D. H. McKnight. 2006. Perceived information quality in data exchanges: Effects on risk, trust, and intention to use. *Inf. Syst. Res.* **17**(4): 332–351.

Okano, T., D. Kitakoshi, M. Suzuki. 2013. A preliminary study on preventive care system based on game playing with communication robots. Proceeding of the 2013 Conference on Technologies and Applications of Artificial Intelligence. Washington, USA.

Ongena, Y. P., M. Haan, D. Yakar, T. C. Kwee. 2020. Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire. *Eur. Radiol.* **v**(2): 1033–1040.

Oscar, S. 2018. The discourse is unhinged': how the media gets ai alarmingly wrong. Available at: https://www.theguardian.com/technology/2018/jul/25/ai-artificial-intelligence-social-media-bots-wrong (accessed date July 30, 2018)

Özer, Ö., Y. Zheng, Y. Ren. 2014. Trust, trustworthiness, and information sharing in supply chains bridging China and the United States. *Manag. Sci.* **60**(10): 2435–2460.

Palanica, A., P. Flaschner, A. Thommandram, M. Li, Y. Fossat. 2019. Physicians' perceptions of chatbots in health care: cross-sectional web-based survey. *J. Med. Internet Res.* **21**(4): e12887.

Patti, C. H., C. H. Chen. 2009. Types of word-of-mouth messages: information search and credence-based services. *J. Promot. Manag.* **15**(3): 357–381.

Pi, S. M., H. L. Liao, H. M. Chen. 2012. Factors that affect consumers' trust and continuous adoption of online financial services. *Int J Bus Manag.* **7**(9): 108.

Pouladzadeh, P., P. Kuhad, S. V. B. Peddi, A. Yassine, S. Shirmohammadi. 2014. Mobile cloud based food calorie measurement. Proceeding of 2014 IEEE International Conference on Multimedia and Expo Workshops July 14-18, Chengdu, China

Preil, D., and Krapp, M. 2021. Artificial intelligence-based inventory management: A Monte Carlo tree search approach. *Ann. Oper. Res.* **60**: 881-936.

Punyatoya, P. 2019. Effects of cognitive and affective trust on online customer behavior. *Mark. Intell. Plan* **37**(1): 66-79.

Rai, A., P. Constantinides, S. Sarker. 2019. Editor's comments: next-generation digital platforms: toward human–AI hybrids. *MIS Q.* **43**(1): 3-15.

Reichheld, F. F., W. E. Sasser. 1990. Zero defections: quality comes to services. *Harv. Bus. Rev.* **68**(5): 105–111.

Rempel, J., J. Holmes, M. Zanna. 1985. Trust in close relationships. *J. Pers. Soc. Psychol.* **49**: 95–112.

Rosenbaum, P. R., D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**(1): 41–55.

Sallis, R. 2015. Exercise is medicine: a call to action for physicians to assess and prescribe exercise. *Phys. Sportsmed.* Taylor & Francis, 22–26.

Schifter, D., I. Ajzen. 1985. Intention, perceived control, and weight loss: an application of the theory of planned behavior. *J. Pers. Soc. Psychol.* **49**: 843–51.

Schierberl, L. 2019. New invoca research conducted by the harris poll, Available at: https://blog.invoca.com/new-research-from-invoca-and-the-harris-poll/ (accessed date June 30, 2019)

Sheeran, P. 2002. Intention—behavior relations: a conceptual and empirical review. *Eur. Rev. Soc. Psychol.* **12**(1): 1–36.

Sheeran, P., S. Orbell. 1999. Implementation intentions and repeated behaviour: Augmenting the predictive validity of the theory of planned behaviour. *Eur. J. Soc. Psychol.* **29**(2–3): 349–369.

Slater, S. F., J. J. Mohr. 2006. Successful development and commercialization of technological innovation: insights based on strategy type. *J. Prod. Innov. Manag.* **23**(1): 26–33.

Song, H., A. L. Tucker, K. L. Murrell, D. R. Vinson. 2018. Closing the productivity gap: improving worker productivity through public relative performance feedback and validation of best practices. *Manag. Sci.* **64**(6): 2628–2649.

Stein, N., K. Brooks. 2017. A fully automated conversational artificial intelligence for weight loss: longitudinal observational study among overweight and obese adults. *JMIR Diabetes* **2**(2): e8590.

Tapus, A., M. J. Mataric, B. Scassellati. 2007. Socially assistive robotics [Grand Challenges of Robotics]. *IEEE Robot. Autom. Mag.* **14**(1): 35–42.

Triandis, H. C. 1979. Values, attitudes, and interpersonal behavior. *Nebr. Symp. Motiv.* **27**: 195–259.

Turkle, S. 2005. *The Second Self, Twentieth Anniversary Edition: Computers and the Human Spirit*. MIT Press.

U.S. Preventive Services Task Force. 1996. Guide to Clinical Preventive Services, 2nd ed. Williams & Wilkins, Maryland, USA.

Venkatesh, V., M. G. Morris, G. B. Davis, F. D. Davis. 2003. User acceptance of information technology: Toward a unified view. *MIS Q.* 425–478.

Venkatesh, V., J. Y. L. Thong, F. K. Y. Chan, P. J. H. Hu. 2016. Managing citizens' uncertainty in e-government services: the mediating and moderating roles of transparency and trust. *Inf. Syst. Res.* **27**(1): 87–111.

Wallace, N., and Castro, D. 2018. The impact of the eu's new data protection regulation on AI. Available at: http://www2.datainnovation.org/2018-impact-gdpr-ai.pdf (accessed date: October 16, 2021)

Wilson, H.J. and Daugherty, P.R., 2018. Collaborative intelligence: humans and AI are joining forces. *Harv. Bus. Rev.* **96**(4):114-123.

World Health Organization. 2016. Health workforce requirements for universal health coverage and the sustainable development goals. Available at https://www.who.int/hrh/resources/health-observer17/en/ (accessed date October 27, 2021).

Yokotani, K., G. Takagi, K. Wakashima. 2018. Advantages of virtual agents over clinical psychologists during comprehensive mental health interviews using a mixed methods design. *Comput. Hum. Behav*. **85**: 135–145.

Yu, K., L. Beam, S. Kohane. 2018. Artificial intelligence in healthcare. *Nat. Biomed. Eng.* **2**(10): 719–731.

Zhang, Y., P. Szolovits. 2008. Patient-specific learning in real time for adaptive monitoring in critical care. *J. Biomed. Inform.* **41**(3): 452–460.

Zhou, M., Y. Fukuoka, Y. Mintz, K. Goldberg, P. Kaminsky, E. Flowers, A. Aswani. 2018. Evaluating machine learning–based automated personalized daily step goals delivered through a mobile phone app: randomized controlled trial. *JMIR MHealth UHealth* **6**(1): e9117.

ONLINE APPENDIX

Rationally Trust, but Emotionally? The Roles of Cognitive and Affective Trust in Laypeople's Acceptance of AI for Preventive Care Operations

SURVEY ITEMS

The survey items for the three theoretical constructs of trust, i.e., ability, benevolence, and integrity, with the supporting literature that proposed the items are provided in Table A1. While we adopted all the survey items from previous studies and revised them to fit into our context, we discarded one item measuring integrity: "I expect that AI (health expert) will keep promises they make." This is because we did not demonstrate any benefits or expected outcomes in our interventions (e.g., "If you achieve the goal, you can lose weight").

Table A1. Survey Constructs and Items for Trust

Construct	Definition	Item No.	Supporting Literature		Item
Cognitive Trust: Ability	Individual perceptions about AI or health expert's knowledge and competence in providing quality health recommendations	AB1	Amazon.com is competent	Gefen and Straub (2004), Gefen (2002)	AI (health expert) is competent in providing health recommendations
		AB2	Amazon.com understands the market it works in		AI (health expert) understands the needs of users
		AB3	Amazon.com knows about books		AI (health expert) knows about health recommendations
		AB4	Amazon.com knows how to provide excellent service		AI (health expert) knows how to provide effective health recommendations
Affective Trust: Benevolence	Individual perceptions about how much AI or health expert cares about its users	BE1	I expect I can count on Amazon.com to consider how its actions affect me		AI (health expert) knows how its health recommendation affects users
		BE2	I expect that Amazon.com has good intentions toward me		AI (health expert) has good intentions toward me
		BE3	I expect that Amazon.com's intentions are benevolent		AI (health expert)'s intentions are benevolent
		BE4	I expect that Amazon.com puts customers' interests before its own		AI (health expert) puts users' interests before their own
		BE5	I expect that Amazon.com is well-meaning		AI (health expert) is well-meaning
Affective Trust: Integrity	Individual perceptions about the AI or health expert's good faith and honesty	IN1	Promises made by Amazon.com are likely to be reliable		Health recommendations made by AI (health expert) are reliable
		IN2	I do not doubt the honesty of Amazon.com		AI (health expert) is honest
		IN3	I expect that advice given by Amazon.com is their best judgment		Health recommendations given by AI (health expert) are its best judgment
		IN4	I can count on Amazon.com to be sincere		AI (health expert) is sincere

QUALITY OF SURVEY

For our survey regarding the four interventions (i.e., AI-, Human-, AI-Human, and AI-transparency interventions), we collected 213 responses while discarding seven low-quality samples. Among the 213 respondents, 51.31% of the respondents are female, and the average age is 50.07. Their average weight and height are 65.22kg and 163.23cm, respectively. The demographic profile of our survey respondents is largely consistent with the experimental subjects. In addition, for all the constructs, Cronbach's alphas are greater than the suggested cutoff value of 0.70 (Hair et al. 2011), and factor loadings are greater than the cutoff value of 0.50 (Truong and McColl 2011, Hulland 1999, Fornell and Larcker 1981). Further, their average variance extracted (AVE) values are above 0.50 (Fornell and Larcker 1981), and their squared roots of AVEs are greater than the correlation between themselves and the other constructs (Hair et al. 2016). These results collectively ensure the quality of our survey.

PROPENSITY SCORE MATCHING

In our main experiment, we observed that *Previous Steps* is not balanced across the three interventions groups ($p < 0.1$). Such a difference, rather than the treatments, could have led to the different acceptance of AI- and Human-interventions. To address the potential selection bias, we employed the propensity score matching (PSM) method (Rosenbaum and Rubin 1983). Please note that we have two different treatment groups (i.e., AI- and Human-interventions) in the main experiment. Following the approaches of previous studies with more than one treatment group (e.g., Liu et al. 2016), we first matched users in the Neutral-interventions group with those in the AI-interventions group, thereby generating the matched pairs of (1) the Neutral- and AI-interventions groups. We then applied the same procedure to the AI- and Human-interventions groups to generate the matched pairs of (2) the AI- and Human-interventions groups. For the matching procedure, we estimated the propensity scores or the probabilities for users to be assigned to each group using logistic regression, which included all the covariates: *Age*, *Gender*, *Height*, *Weight*, *Previous Steps*, *App Proficiency*, and *Location*. Given the estimated propensity score of each user, the nearest neighbor matching without replacement was employed with a caliper width of 0.01. To identify the statistical differences in the covariates between the matched groups, we conducted a series of *t*-tests and found that all the covariates are statistically similar after matching ($p > 0.1$) (see Tables A2 and A3 for the results).

We replicated the main analysis using the two sets of matched samples. The first and second panels of Table A4 show the results of (1) the Neutral- and AI-interventions groups and (2) the AI- and Human-interventions groups, respectively. Please note that the Neutral- and AI-interventions groups were used as the baseline groups for each analysis, respectively. As a result, we find that AI-interventions are significantly more effective than Neutral-interventions (see the first panel of Table A4), while their effectiveness is significantly lower than that of Human-interventions (see the second panel of Table A4). In other words, the main results are consistent even after considering the potential selection bias.

Table A2. Covariate Balance before and after Matching Neutral- and AI-interventions Groups

Variables	Before Matching					After Matching				
	Neutral-Interventions		AI-Interventions		*t* stat.	Neutral-Interventions		AI-Interventions		*t* stat.
	Mean	S.D	Mean	S.D		Mean	S.D	Mean	S.D	
Female	0.54	0.49	0.53	0.49	-0.75	0.54	0.49	0.55	0.49	0.33
Age	51.05	13.90	50.97	13.74	-0.53	51.54	13.25	51.62	13.46	0.20
Weight	64.52	17.15	65.59	17.27	0.81	64.35	13.53	64.36	13.63	0.02
Height	164.67	10.15	164.51	10.95	-0.20	164.10	10.45	164.08	10.58	-0.08
Previous Steps	49359.24	31688.17	49758.91	31989.82	0.67	48282.81	30637.74	48496.11	30817.15	0.25
App Proficiency	3.373	9.3527	3.585	9.3210	0.88	3.48	9.371	3.46	9.287	0.31

Note: Mean values of the fifteen Location dummy variables are also statistically similar after matching ($p > 0.1$). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3. Covariate Balance before and after Matching AI- and Human-interventions Groups

Variables	Before Matching					After Matching				
	AI-Interventions		Human-Interventions		*t* stat.	AI-Interventions		Human-Interventions		*t* stat.
	Mean	S.D	Mean	S.D		Mean	S.D	Mean	S.D	
Female	0.53	0.49	0.54	0.49	0.85	0.53	0.49	0.53	0.49	-0.03
Age	50.97	13.74	51.04	13.89	0.18	51.06	13.55	51.18	13.56	0.11
Weight	65.59	17.27	64.54	17.61	-0.60	65.19	12.12	65.25	12.77	0.15
Height	164.51	10.95	164.08	10.94	-0.52	164.52	9.50	164.42	9.38	-0.37
Previous Steps	49758.91	31989.82	50177.93	37317.04	1.46**	49638.92	31824.73	49454.38	35810.94	-0.20
App Proficiency	3.585	9.3210	3.507	9.293	0.93	3.58	9.65	3.68	9.61	0.67

Note: Mean values of the fifteen Location dummy variables are also statistically similar after matching ($p > 0.1$). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A4. Results for Matched Samples

Matched Groups	**(1) Neutral- and AI-interventions**		**(2) AI- and Human-interventions**	
Variables	**Acceptance**	**Achievement**	**Acceptance**	**Achievement**
AI	0.577***	0.627***		
	(0.0756)	(0.0999)		
Human			0.227***	0.289***
			(0.0654)	(0.0824)
Constant	-2.897***	-2.465*	-1.802**	-2.009**
	(0.688)	(1.382)	(0.736)	(0.887)
Controls	Yes	Yes	Yes	Yes
Observations	5,781	5,781	5,760	5,760
Log-likelihood	-2381.83	-1585.908	-2910.59	-2066.73

Note: The baseline groups for Panels 1 and 2 are the Neutral- and AI-interventions groups, respectively. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TWO-STAGE PROCESS OF ACCEPTANCE AND ACHIEVEMENT

In our experiment, a user made the two decisions in sequence: (1) whether to accept an intervention (i.e., *Acceptance*) and (2) whether to complete it (i.e., *Achievement*). Thus, the outcome of the second decision could only be observed when a user accepted an assigned intervention, thereby raising concerns regarding such a two-stage process. To alleviate this concern, we first replicated the analyses of *Achievement* only for users who accepted their interventions. This can capture the *efficacy* of different interventions, i.e., the proportion of individuals who achieve desired health behavior to those who accept the given intervention, which is one of the important practical measures of preventive health intervention effectiveness (Halpern et al. 2015). Table A5 shows the results, which are qualitatively consistent with the main findings. Thus, we can generalize our findings toward different performance measures of preventive health interventions, which improves the practical value of this study.

Table A5. Results Only for Accepted Users

(1) Main Interventions		(2) Additional Interventions	
Variables	**Achievement**	**Variables**	**Achievement**
AI	0.246***	AI-Human	0.791***
	(0.076)		(0.0860)
Human	0.792***	AI-Transparency	0.387***
	(0.018)		(0.0930)
Constant	-4.127***	Constant	-0.937***
	(1.069)		(0.222)
Controls	Y	Controls	Y
Observations	1,572	Observations	2,044
Log-likelihood	-505.847	Log-likelihood	-1012.5

Note: Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In addition, we adopted Heckman's selection model (Heckman 1976) to address a potential concern about the two-stage process of the outcomes. For the analysis of the main interventions, we started with a Probit model for the decision in the second stage, i.e., *Achievement*:

$$Achievement_i^* = \alpha_1 * AI_i + \alpha_2 * Human_i + \beta * X_i + \varepsilon_i,$$
$$Achievement_i = \begin{cases} 1, & Achievement_i^* > 0 \\ 0, & Otherwise \end{cases}, \tag{A1}$$

where *i* represents each user, and $Achievement_i^*$ indicates a latent variable that determines the likelihood of *Achievement* for individual i. $Achievement_i^*$ depends on AI_i and $Human_i$, i.e., binary variables indicating whether user *i* received AI- or Human-interventions, respectively; vector of observed individual characteristics $X_{i;}$, and random error ε_i. Please note that the Neutral-interventions group was used as the baseline group. We included the same covariates in X_i as the main analysis.

Given *Achievement*, which can be observed only when users tap the "Yes" button in the assigned intervention (i.e., *Acceptance*= 1), we incorporated the first-stage decision, i.e., *Acceptance*, in the analysis by using Heckman's selection model (Heckman 1976).

$$Acceptance_i^* = \gamma_1 * AI_i + \gamma_2 * Human_i + \delta * X_i + \phi Z_i + u_i$$

$$Acceptance_i = \begin{cases} 1, & Acceptance_i^* > 0 \\ 0, & Otherwise \end{cases}, \qquad \text{(A2)}$$

$$Achievement_i \text{ is observed only if } Acceptance_i = 1,$$

where $Acceptance_i^*$ is an unobserved latent variable that determines the likelihood of *Acceptance* for individual i, and it depends on AI_i, $Human_i$, X_i, vector of selection variables Z_i, and random error u_i. Note that in the first-stage model (i.e., Equation A2), we used binary variable *Metropolitan* indicating whether a user lives in a metropolitan area or not as selection variable Z_i. The rationale is that users who live in metropolitan areas would have a greater propensity to tap the "Yes" button in mobile interventions, compared with those who live in rural areas, due to their higher proficiency in and familiarity with mobile services and technologies. The key feature of Heckman's selection model is a correlation between the two unobserved error terms, each of which affects *Acceptance* and *Achievement*, i.e., $Corr\ (\varepsilon, u) = p$. We estimate the models with the maximum likelihood estimator.

As can be seen in Table A6, *Metropolitan* has a significant and positive effect in the first stage of the decision ($p < 0.1$). In addition, we found that AI-interventions are still significantly more and less effective than Neutral-interventions and Human-interventions, respectively, in terms of both *Acceptance* and *Achievement*. We also replicated the same analysis for the additional interventions and found that the results remained consistent (see Table A7); AI-Human and AI-Transparency interventions are significantly more effective than AI-interventions. Thus, the consistent results further alleviate the concern for the two-stage structure of the outcome variables.

Table A6. Results of Two-Stage Selection Model for the Main Interventions

Variables	(1) Acceptance	(2) Achievement
AI	0.132***	0.320**
	(0.036)	(0.136)
Human	0.455***	0.733***
	(0.039)	(0.121)
Metropolitan	0.065*	
	(0.039)	
Constant	–1.617***	–4.614***
	(0.334)	(1.214)
Controls	Y	Y
Observations	9,000	1,572
Log-likelihood	–4,594.409	

Note: The baseline group is the Neutral-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7. Results of Two-Stage Selection Model for the Additional Interventions

Variables	(1) Acceptance	(2) Achievement
AI-Human	0.287***	0.186***
	(0.0361)	(0.0383)
AI-Transparency	0.120***	0.517***
	(0.0368)	(0.0380)
Metropolitan	0.0227*	
	(0.014)	
Constant	-1.2609***	-0.9059***
	(0.3331)	(0.3106)
Controls	Y	Y
Observations	9,000	2,044
Log-likelihood	-5813.86	

Note: The baseline group is the AI-interventions group. Controls include demographic (i.e., Age, Gender), physical (i.e., Height, Weight), behavioral (i.e., Previous Steps), technical (i.e., App Proficiency), and geographical (i.e., Location) variables. Robust standard errors are in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

GAP BETWEEN PREVIOUS BEHAVIOR AND SUGGESTED GOAL

The difference between the previous number of steps and the suggested step goal could have a direct or moderating impact on users' responsiveness to different interventions (i.e., *Acceptance* and *Achievement*). To investigate this possibility, we replicated the main analyses with an additional control variable *Gap* (> 0), which indicates the number of steps walked in the past seven days subjected from the suggested step goal. Tables A8 and A9 show the results of the main and additional experiments, respectively. The first and third columns of Tables A8 and A9 show that *Gap* has an insignificant direct impact on both dependent variables (i.e., *Acceptance* and *Achievement*). We additionally included its interactions with different interventions, and the results also demonstrated their insignificant impacts on both dependent variables (see the second and fourth columns of Tables A8 and A9). These results imply that individuals' different responsiveness to various interventions is not driven by the difference between the number of steps walked in the past seven days and the suggested step goal.

Table A8. Results of Gap for Main Interventions

Variables	(1) Acceptance	(2) Acceptance	(3) Achievement	(4) Achievement
AI	0.582***	0.550***	0.623***	0.605***
	(0.0832)	(0.0975)	(0.112)	(0.126)
Human	0.815***	0.841***	0.902***	0.905***
	(0.0782)	(0.0911)	(0.102)	(0.116)
Gap	-0.0569	-0.0540	-0.123	-0.143
	(0.107)	(0.152)	(0.142)	(0.214)
AI * Gap		0.102		0.0682
		(0.161)		(0.225)
Human * Gap		-0.0907		-0.0111
		(0.159)		(0.219)
Constant	-2.733***	-2.731***	-2.807***	–2.753***
	(0.574)	(0.573)	(0.575)	(0.569)
Controls	Y	Y	Y	Y
Observations	9,000	9,000	9,000	9,000
Log-likelihood	-4088.7843	-4088.9607	-2824.661	-2822.5972

Note: The baseline group is the *Neutral-interventions* group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A9. Results of Gap for Additional Interventions

Variables	(1) Acceptance	(2) Acceptance	(3) Achievement	(4) Achievement
AI-Human	0.498***	0.496***	0.718***	0.708***
	(0.0625)	(0.0632)	(0.0764)	(0.0773)
AI-Transparency	0.213***	0.209***	0.260***	0.259***
	(0.0645)	(0.0652)	(0.0816)	(0.0824)
Gap	-0.0250	-0.112	-0.121	-0.284
	(0.182)	(0.347)	(0.219)	(0.484)
AI-Human * Gap		0.0783		0.466
		(0.423)		(0.548)
AI-Transparency * Gap		0.155		0.113
		(0.431)		(0.596)
Constant	-1.497***	-1.499***	-1.637***	-1.624***
	(0.571)	(0.572)	(0.609)	(0.609)
Controls	Y	Y	Y	Y
Observations	9,000	9,000	9,000	9,000
Log-likelihood	-4779.998	-4088.96	-3583.03	-2822.59

Note: The baseline group is the *Neutral-interventions* group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

SOFTER REQUEST FOR HUMAN-INTERVENTIONS

The results of Human-interventions, which were delivered by “health experts,” could be different if we change the last sentence of Human-interventions (i.e., “Would you like to participate?”) to a softer request. This is because the last sentence of Human-interventions could unconsciously highlight the authority of health experts or imbue users with some social obligation, thereby leading to a greater acceptance of Human-interventions. First of all, please note that the last sentence is an interrogative sentence that asks a question without intention to convince or impose a social obligation on listeners (Sohn 2001). Thus, the extent to which the last sentence carries some form of authority would be limited.

Nevertheless, we conducted an additional experiment in collaboration with the focal firm to address the concern. Specifically, we identified whether the use of a *softer* request would result in a *lower* acceptance of Human-interventions. To rule out such a possibility, we revised the last sentence of Human-interventions from "Would you like to participate?" to "Come join us for your health." We then randomly sent out each of the original and revised Human-interventions (i.e., with "Would you like to participate?" and "Come join us for your health," respectively) to 50 unique users of the focal app. If the alternative explanation holds true, the acceptance rate of the revised Human-interventions with the softer request should be lower than that of the original Human-interventions due to their reduced authority or social obligation.

However, the results reveal that the difference in the acceptance rate between the original and revised Human-interventions is statistically insignificant (*t*-test, $p > 0.1$) (see Table A10). Moreover, the acceptance rate of the revised Human-interventions is greater than that of the original Human-interventions. These results are in contrast with the expected outcome under the alternative scenario, thereby alleviating the concern about the authority or social obligation regarding Human-interventions.

Table A10. Result of Human-Interventions with Different Tones

Group	Intervention Message	Sample	Acceptance	Difference
Original Human-interventions	Health expert recommends that you walk [*personalized step goal*] in the next seven days. Would you like to participate?	50	5	*t*-test: $p > 0.1$
Revised Human-interventions	Health expert recommends that you walk [*personalized step goal*] in the next seven days. Come join us for your health.	50	8	

ALTERNATIVE CONTROL GROUP

The use of an additional question mark only in Neutral-interventions would inadvertently *enhance* or *reduce* individuals' responsiveness only to Neutral-interventions, compared to AI- and Human-interventions. To alleviate the concern regarding our control group, we constituted an alternative control group by randomly selecting 3,000 unique users who did not receive any interventions. We then used the same AI algorithm to calculate this *No-Intervention* group's latent step goal and measured *Achievement* as whether each user in this group walked more than their latent step goal or not during the treatment period. Please note that we cannot measure this group's *Acceptance* since they did not receive any interventions to accept. We replicated the main analysis using this *No-Intervention* group as an alternative control group (see Table A11). The results are consistent with the main results.

Table A11. Result of Using No-Intervention Group as Alternative Baseline

Variables	Achievement
AI	1.687***
	(0.153)
Human	2.173***
	(0.151)
Constant	-4.643***
	(0.799)
Controls	Y
Observations	9,000
Log-likelihood	-2410.34

Note: The baseline group is the *No-interventions* group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

RESULTS FOR READERS OF INTERVENTION MESSAGES

The distribution of users who did not read the intervention messages could be unbalanced across the different intervention groups, thereby affecting the results. To identify how many of the users did not read the message and whether such non-readers affect the results, we asked the focal firm to provide us with additional data on which button users tapped on the pop-up window. When users received one of the interventions, a pop-up window with the intervention message was opened (see Figure 1 in the manuscript). This pop-up window will not be closed until users tap the Yes, No, or Close (X) button. While we cannot exactly identify whether users read the intervention message or not, it can be generally assumed that users who tapped the Close (X) button did not read the message (i.e., non-readers), while those who tapped the Yes or No button read it (i.e., readers).

As can be seen in Table A12, based on our assumption, only around 2% of our sample did not read the intervention messages (i.e., tapped Close (X) button). We conducted an analysis of variance (ANOVA) to compare the number of non-readers across the different intervention groups. The results indicate that the number of non-readers is balanced and statistically similar across the different groups ($p > 0.1$). In addition, we replicated the main analyses only for readers. Tables A13 and A14 show the results, which are qualitatively consistent with the main results. These results collectively address the concern for non-readers.

Table A12. Distribution of Non-readers across Different Groups

Group	Obs.	Proportion of Non-Readers (%)
Neutral	59	1.96
AI	47	1.57
Human	56	1.87
AI-Human	74	2.47
AI-Transparency	63	2.10

Table A13. Results Without Non-Reader for the Main Interventions

Variables	(1) Acceptance	(2) Achievement
AI	0.590***	0.621***
	(0.074)	(0.098)
Human	0.820***	0.907***
	(0.072)	(0.095)
Constant	–2.787***	–3.229***
	(0.652)	(0.859)
Controls	Y	Y
Observations	8,838	8,838
Log-likelihood	–4,079.09	–2,814.81

Note: The baseline group is the Neutral-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A14. Results Without Non-Reader for the Additional Interventions

Variables	(1) Acceptance	(2) Achievement
AI-Human	0.491***	0.713***
	(0.0627)	(0.0766)
AI-Transparency	0.206***	0.252***
	(0.0646)	(0.0818)
Constant	-1.478***	-1.638***
	(0.573)	(0.615)
Controls	Y	Y
Observations	8,816	8,816
Log-likelihood	-4751.24	-3558.784

Note: The baseline group is the AI-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TREATMENT EFFECTS FOR PEOPLE IN DIFFERENT AGE GROUPS

Given that the average age of our sample is around 50, there can be a concern about whether the results could be generalized to different age groups (e.g., millennials, the elderly). For example, the elderly tend to have low technology acceptance and typically do not trust technology as young adults do. This could lead to biased estimates of the relative acceptance of AI- and Human-interventions. To address the concern, we classified users into the three age groups, i.e., under 30 (i.e., millennials), 30–59, and over 60 (i.e., the elderly), and replicated our main analyses for users in each age group. As can be seen in Tables A15 - A18, the results remain qualitatively consistent, thereby addressing the generalizability concern.

Table A15. Results of Acceptance by Age Groups for the Main Interventions

Variables		**Acceptance**	
Groups	**(1) Under 30**	**(2) 30–59**	**(3) Over 60**
AI	2.296***	1.589***	0.671***
	(0.976)	(0.388)	(0.153)
Human	2.670***	1.836***	0.921***
	(0.972)	(0.386)	(0.149)
Constant	–1.359*	–3.084***	–7.558***
	(0.799)	(1.063)	(1.564)
Controls	Y	Y	Y
Observations	801	6,031	2,168
Log-likelihood	–716.550	–2,902.003	–973.349

Note: The baseline group is the Neutral-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A16. Results of Achievement by Age Groups for the Main Interventions

Variables		**Achievement**	
Groups	**(1) Under 30**	**(2) 30–59**	**(3) Over 60**
AI	3.009***	1.962***	1.575***
	(1.053)	(0.190)	(0.275)
Human	3.637***	2.214***	1.907***
	(1.036)	(0.188)	(0.271)
Constant	–5.826***	–5.011***	–3.423**
	(1.955)	(1.513)	(1.434)
Controls	Y	Y	Y
Observations	801	6,031	2,168
Log-likelihood	–155.28	–1,638.22	–602.31

Note: The baseline group is the Neutral-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A17. Results of Acceptance by Age Groups for the Additional Interventions

Variables		**Acceptance**	
Groups	**(1) Under 30**	**(2) 30–59**	**(3) Over 60**
AI-Human	0.727***	0.359**	0.640***
	(0.0915)	(0.130)	(0.103)
AI-Transparency	0.240**	0.224*	0.394***
	(0.111)	(0.123)	(0.107)
Constant	-2.394**	-0.442	-0.336*
	(1.042)	(1.237)	(0.196)
Controls	Y	Y	Y
Observations	1,092	5,392	2,516
Log-likelihood	-247.233	-639.20	-2691.84

Note: The baseline group is the AI-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A18. Results of Achievement by Age Groups for the Additional Interventions

Variables	Achievement		
Groups	**(1) Under 30**	**(2) 30–59**	**(3) Over 60**
AI-Human	1.123***	0.499***	0.446***
	(0.026)	(0.0765)	(0.073)
AI-Transparency	0.495**	0.197**	0.140*
	(0.134)	(0.0793)	(0.076)
Constant	-1.807*	-2.410*	-1.268
	(0.991)	(1.267)	(0.920)
Controls	Y	Y	Y
Observations	1,092	5,392	2,516
Log-likelihood	-389.89	-869.83	-2405.49

Note: The baseline group is the AI-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TREATMENT EFFECTS FOR PEOPLE IN NEED OF EXERCISE

Researchers and practitioners have argued that walking 10,000 steps a day is “a reasonable target for healthy adults (Tudor-Locke et al. 2011)” to meet the *minimum* physical activity guideline (e.g., Le-Masurier et al. 2003). In this regard, 21% of our sample already achieved this modest goal before the experiment. To identify whether our results are biased by relatively healthy people, we replicated the main analyses for users whose *Previous Steps* are less than 70,000 (i.e., 10,000 steps a day). Tables A19 and A20 demonstrate that our results remain consistent, thereby alleviating the concern for generalizability toward people in need of exercise.

Table A19. Results for Users in Need of Exercise for the Main Interventions

Variables	(1) Acceptance	(2) Achievement
AI	0.573***	0.602***
	(0.082)	(0.110)
Human	0.804***	0.852***
	(0.080)	(0.106)
Constant	-2.5993***	–3.529***
	(0.6192)	(0.986)
Controls	Y	Y
Observations	7,159	7,159
Log-likelihood	–3,287.09	–2,244.13

Note: The baseline group is the Neutral-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A20. Results for Users in Need of Exercise for the Additional Interventions

Variables	(1) Acceptance	(2) Achievement
AI-Human	0.595***	0.640***
	(0.0872)	(0.103)
AI-Transparency	0.257***	0.394***
	(0.0918)	(0.107)
Constant	-1.180*	-1.268
	(0.692)	(0.920)
Controls	Y	Y
Observations	7,106	7,106
Log-likelihood	-2040.96	-2691.884

Note: The baseline group is the Neutral-interventions group. Controls include demographic (i.e., *Age*, *Gender*), physical (i.e., *Height*, *Weight*), behavioral (i.e., *Previous Steps*), technical (i.e., *App Proficiency*), and geographical (i.e., *Location*) variables. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

RULING OUT ALTERNATIVE MECHANISM FOR DIAGNOSTIC CARE

In the context of *diagnostic* care, previous studies suggest that individuals prefer humans to AI when receiving a diagnostic test because of the concern that AI neglects their unique, personal circumstances (i.e., uniqueness neglect) (Longoni et al. 2019). Accordingly, including more personalization features is expected to improve users' acceptance of AI for diagnostic care. If this explanation also holds true in the context of preventive care, the lower acceptance of AI-intervention would be due to uniqueness neglect, instead of the lack of affective trust. To examine the validity of this alternative mechanism in the context of preventive care, we included the three measurement items for *uniqueness neglect* suggested by Longoni et al. (2019) in our survey of the AI- and Human-interventions groups. The survey items for *uniqueness neglect* with the supporting literature that proposed the items are provided in Table A21.

Table A21. Survey Constructs and Items for Uniqueness Neglect

Construct	Definition	Item No.	Supporting Literature		Item
Uniqueness Neglect	Individual perceptions that their characteristics are unique and different from those of others	UN1	The recommender would not recognize the uniqueness of your medical condition	Longoni et al. (2019)	AI (health expert) recognizes the uniqueness of your health condition
		UN2	The recommender would not consider your unique circumstances		AI (health expert) considers your unique circumstances
		UN3	The recommender would not tailor the recommendation to your unique case		AI (health expert) tailors the health recommendation to your unique case

For the analysis, we compare the mean values of uniqueness neglect for AI- and Human-interventions groups. First, the results show that Human-interventions, which have a greater acceptance rate than AI-interventions, have lower uniqueness neglect (2.572) compared to AI-interventions (2.741), while the difference is statistically insignificant (*t*-test, $p > 0.1$). This implies that individuals' preference for Human-interventions over AI-interventions cannot be explained by the uniqueness neglect of AI. One possible explanation for the result is that AI for preventive care operations already provides sufficient *personalization* features, such as "*personalized step goal*" in our intervention messages. Thus, in the context of preventive care, we can rule out the alternative mechanism suggested for AI for diagnostic care. The result further improves the theoretical validity of the study.

FULL RESULT TABLES

Table A22. Results of the Main Interventions

Variables	(1) Acceptance	(2) Achievement
AI	0.589***	0.609***
	(0.0743)	(0.0985)
Human	0.819***	0.896***
	(0.0724)	(0.0946)
Age	0.00134	0.00360
	(0.00218)	(0.00263)
Gender	0.125*	0.00118
	(0.0704)	(0.0926)
Weight	0.138	0.311
	(0.222)	(0.239)
Height	0.0875	0.0853
	(0.195)	(0.239)
Previous Steps	0.00231	0.0553
	(0.139)	(0.179)
App Proficiency	0.00120	0.00619
	(0.00307)	(0.128)
Location1	0.00372	0.0443
	(0.0778)	(0.0913)
Location2	-0.0296	0.0435
	(0.0655)	(0.0769)
Location3	0.0404	0.0452
	(0.0669)	(0.0790)
Location4	-0.361	0.0794
	(0.380)	(0.401)
Location5	-0.0261	-0.211
	(0.129)	(0.164)
Location6	-0.179	0.177
	(0.243)	(0.242)
Location7	-0.0589	0.360**
	(0.131)	(0.136)
Location8	0.0329	-0.0616
	(0.0600)	(0.0719)
Location9	0.143	0.237
	(0.163)	(0.237)

Location10	0.260	-0.155
	(0.175)	(0.235)
Location11	0.00722	0.0677
	(0.0718)	(0.0835)
Location12	-0.317***	0.121
	(0.120)	(0.124)
Location13	-0.0499	0.0516
	(0.0945)	(0.107)
Location14	0.0214	0.128
	(0.122)	(0.139)
Constant	–2.753***	–3.662***
	(0.569)	(0.856)
Observations	9,000	9,000
Log-likelihood	–4,089.02	–2,822.92

Note: The baseline group is the *Neutral-interventions* group. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A23. Results of the Additional Interventions

Variables	(1) Acceptance	(2) Achievement
AI-Human	0.498***	0.718***
	(0.0625)	(0.0764)
AI-Transparency	0.212***	0.259***
	(0.0644)	(0.0816)
Age	0.000938	0.00339
	(0.00191)	(0.00224)
Gender	0.0306	0.110*
	(0.0645)	(0.0645)
Weight	0.0660	0.0781
	(0.119)	(0.154)
Height	0.137	0.181
	(0.117)	(0.140)
Previous Steps	0.00677	0.00493
	(0.107)	(0.00346)
App Proficiency	0.00200	0.00164
	(0.00276)	(0.00309)
Location1	-0.142*	-0.174**
	(0.0773)	(0.0865)
Location2	0.00399	-0.122*
	(0.0629)	(0.0709)
Location3	-0.108*	-0.141*
	(0.0647)	(0.0725)
Location4	-0.325	0.111
	(0.303)	(0.282)
Location5	0.0528	-0.0498
	(0.115)	(0.131)
Location6	-0.288	-0.344
	(0.234)	(0.279)
Location7	-0.0811	0.180
	(0.129)	(0.135)

Location8	-0.0363	-0.149**
	(0.0578)	(0.0650)
Location9	-0.207	0.0909
	(0.457)	(0.470)
Location10	-0.0806	-0.226
	(0.165)	(0.196)
Location11	-0.0802	-0.101
	(0.0680)	(0.0762)
Location12	-0.167	-0.195
	(0.107)	(0.121)
Location13	-0.205**	0.0265
	(0.0904)	(0.0956)
Location14	-0.0149	-0.0669
	(0.107)	(0.120)
Constant	-1.514***	-1.678***
	(0.573)	(0.611)
Observations	9,000	9,000
Log-likelihood	-3582.874	-4779.8793

Note: The baseline group is the AI-interventions group. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

REFERENCES

Fornell, C., D. F. Larcker. 1981. Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res*. **18**(1): 39–50.

Gefen, D. 2002. Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM SIGMIS Database DATABASE Adv. Inf. Syst.* **33**(3): 38–53.

Gefen, D., D. W. Straub. 2004. Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. *Omega* **32**(6): 407–424.

Hair, J. F., C. M. Ringle, M. Sarstedt. 2011. PLS-SEM: Indeed a silver bullet. *J. Mark. Theory Pract.* **19**(2): 139–152.

Halpern, S. D., French, B., Small, D. S., Saulsgiver, K., Harhay, M. O., Audrain-McGovern, J., Volpp, K. G. 2015. Randomized trial of four financial-incentive programs for smoking cessation. *N. Engl. J. Med.* **372**(22): 2108-2117

Hulland, J. 1999. Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strateg. Manag. J.* **20**(2): 195–204.

Le-Masurier, G. C., C. L. Sidman, C. B. Corbin. 2003. Accumulating 10,000 steps: does this meet current physical activity guidelines? *Res. Q. Exerc. Sport* **74**(4): 389–394.

Liu, C.-W., R. Agarwal, G. Gao. 2016. The dark side of positive social influence. Proceeding of the 2016 International Conference on Information Systems, December 11-14, Dublin, Ireland

Longoni, C., A. Bonezzi, C. K. Morewedge. 2019. Resistance to Medical Artificial Intelligence. *J. Consum. Res.* **46**(4): 629–650.

Rosenbaum, P. R., D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**(1): 41–55.

Sohn, H.-M. 2001. *The Korean Language*. Cambridge University Press.

Truong, Y., R. McColl. 2011. Intrinsic motivations, self-esteem, and luxury goods consumption. *J. Retail. Consum. Serv.* **18**(6): 555–561.

Tudor-Locke, C., C. L. Craig, W. J. Brown, S. A. Clemes, K. De Cocker, B. Giles-Corti, Y. Hatano, S. Inoue, S. M. Matsudo, N. Mutrie. 2011. How many steps/day are enough? For adults. *Int. J. Behav. Nutr. Phys. Act.* **8**(1): 1–17.