Is AI Augmenting or Substituting Humans?

An Eye-Tracking Study of Visual Attention Toward Health Application

Piotr Gaczek, Poznan University of Economics and Business, Poland*

<https://orcid.org/0000-0003-0923-7656>

Grzegorz Leszczyński, Poznan University of Economics and Business, Poland

Marek Zieliński, Poznan University of Economics and Business, Poland

<https://orcid.org/0000-0001-5697-3678>

ABSTRACT

In this paper, the authors focus on artificial intelligence as a tangible technology that is designed to sense, comprehend, act, and learn. There are two manifestations of AI in the medical service: an algorithm that analyzes and interprets the test result and a virtual assistant that communicates the result to the patient. The aim of this paper is to consider how AI can substitute a doctor in measuring human health and how the interaction with virtual assistant impacts one's visual attention processes. Theoretically, the article refers to the following research strands: human-computer interaction, technology in services, implementation of AI in the medical sector, and behavioral economy. By conducting an eye-tracking experimental study, it is demonstrated that the perception of medical diagnosis does not differ across experimental groups (human vs. AI).

KEYWORDS

Healthcare, Artificial Intelligence, AI-Human Interaction, Virtual Assistant, Experimental Economics, Eye-tracking, Cognitive Effort

INTRODUCTION

Artificial Intelligence (AI) is being applied to more and more areas of everyday life, including medical diagnostics (McDougall, 2019). In the past two decades, many types of IT-based systems have been broadly utilized in health care in different domains (Rajabion, Shaltooki, Taghikhah, Ghasemi, & Badfar, 2019). Various scenarios have been made possible by rapid advances in information and communications technologies and by the increasing number of smart things (portable devices and sensors) (Kashyap, 2020). The promise of artificial intelligence (AI) in health care offers substantial opportunities to improve patient and clinical team outcomes, reduce costs, and influence the health of the general population. From the patient's point of view, it enables constant monitoring of the body's condition, and analysis of the results through comparison with a vast database containing information no doctor or human analyst would have time to process (Grzywalski, et al., 2019). From the doctor's point of view, it allows them to diagnose patients without having to have contact with them. And the doctor, thanks to AI support, can focus on patients who need direct personal contact (Davenport

DOI: 10.4018/IJTHL.293193

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

& Kalakota, 2019). This raises high expectations for various applications and devices that monitor health using AI, e.g., Patient Monitoring Devices based on Internet of Things (Shah & Chircu, 2018).

In this article, we identify the reactions of patients to such solutions. We are interested in the reaction to the diagnosis issued by AI and the reaction to a message about the result of the diagnosis delivered by artificial intelligence. These are important issues, because it is the users' attitude towards these solutions that will mostly decide about their distribution on the market (Nasirian, Ahmadian, & Lee, 2017). Research on AI is about any device that perceives its environment and takes actions that maximize its chance to successfully achieve its goals (Russell & Norvig, 2009). In this paper, we focus on AI as a tangible technology that is designed to sense, comprehend, act, and learn to be capable of performing tasks that, if performed by a human, would be said to require intelligence (Scherer, 2016). We deal with medical diagnostic services in which two manifestations of AI are present: (1) an algorithm that collects data from the patient and analyzes and interprets the test result and a (2) Virtual Assistant that communicates the result to the patient and manages his or her communication with the doctor.

The subject of our research is user cognitive reactions to the performance of the test by the AI algorithm, and to the Virtual Assistant's communication of this result. This approach is similar to one presented by Wijesinghe, Gamage, Perera, and Chitraranjan (2019) for smart telemedicine system to manage diabetic retinopathy. Empirically, this paper is focused on the Polish health system because it has high hopes for the implementation of telemedicine in order to improve its condition (Domagała & Klich, 2018).

BACKGROUND

AI in Services

One of the fundamentals of service science says that interaction takes a unique role in services because usually the value is delivered to a customer during those interactions (Maglio & Spohrer, 2008). However, it is not always a human actor who interacts with a customer - technology can also play such a role in service delivery. Specifically, technology can play an augmenting or substituting role for a human actor in service delivery (Marinova, De Ruyter, Huang, Meuter, & Challagalla, 2017). It can enhance the human ability to think (e.g., by collecting and analyzing data) and to act (e.g., by communicating). In some settings, technology can interact with customers faster, cheaper, and in a more convenient way than a human, all in the name of service delivery (De Keyser, Köcher, Alkire, Verbeeck, & Kandampully, 2019). It can also be used by a customer who consumes a service, e.g., using a smart band to collect medical data from a person who wears it and sending it to an algorithm that provides real-time analysis and alerts a doctor if the results exceed allowable levels.

According to Marinova et al. (2017), technology can also replace a human actor who delivers a service. A chatbot or Virtual Assistant (VA) can be one example of such substitution as they are made available, thanks to AI-based technology (Kot & Leszczyński, in press). It can communicate with a patient who uses the medical service to present the results of medical tests, comment on those results, and schedule a meeting with a doctor if needed. Bots and VAs can be shaped to appear as human-like, animal-like, cartoon-like, and functional (Fong, Nourbakhsh, & Dautenhahn, 2003) to facilitate natural and effective interaction. Their interactive behavior can be programmed to include an emotional display, gaze, personality, sense of humor, or even gestures and movements (De Keyser et al., 2019). Their interface can be based on text, voice, graphics, touch, or biometrics. Wijesinghe et al. (2019) in a research paper on an Intelligent Diabetic Assistant present such a system which decides the diagnosis and the treatment priority depending upon the observations that have appeared on the screen. Summing up, both sides of health care service interaction can use the technology as a part of the service. It allows clinical services to be delivered to patients remotely, into areas which were previously served by well-trained human clinicians (Aborujilah et al. 2020). Other examples

of such applications in health care are the virtual nurse assistants Sensely (sensely.com) or Angel (careangel.com).

Thus, AI can augment or replace humans in service; it can be utilized by a customer or service company - both scenarios are plausible. It can already be seen in the medical sector, where more and more data on personal health measurements can be collected and analyzed, and doctors are increasingly burdened with more and more cases (Reddy, Fox, & Purohit, 2019). Applying solutions based on AI to support and help is at the very least reasonable in these situations.

Current works on human-AI interaction suggests that in every mode of interaction, an interface plays a vital role, as it facilitates mutual communication (Li, 2015). The communication features of an AI system can influence how humans approach them (Mou & Xu, 2017). A simple AI system does not deliver the capability of self-adaptation that results from an advanced comprehension of human verbal and non-verbal communications and the ability to answer correctly (Fanelli, Gall, Romsdorfer, Weise, & Van Gool, 2010). Advanced, interactive AI systems are able to imitate humans and to mimic human users (Khan & Das, 2018). One of the first studies was conducted in the 1960s at MIT Artificial Intelligence Laboratory (Weizenbaum & McCarthy, 1976). Researchers programmed the AI system to act like a doctor who conducts the therapeutic process to a patient. Despite the simplicity of that system, some patients attributed human-like feelings to it and even wanted to continue the therapy. Today, researchers do not focus so much on tricking AI users (as AI obviously can do it) but on their attitude to AI systems. User satisfaction, sense of personalized service, engagement in conversations, or trust in AI is considered (Przegalińska, Ciechanowski, Stroz, Gloor, & Mazurek, 2019).

AI in the Health Care Sector

In the near future, AI-based applications can potentially be used in a variety of areas related to prevention, treatment, and rehabilitation. AI can support the work of doctors, making it easier to make the right diagnosis or to choose the optimal path of treatment for the patient (Reddy et al., 2019). At the same time, AI indicates an opportunity for the development of personalized medicine. Research on new molecules is moving towards personalized therapies, so the pharmaceutical product can be selected for the patient in terms of their specific genetic profile. So, another area of AI application in the health care sector is in the drug research and development process (Mak & Pichika, 2019). Solutions using intelligent and personal robots are beginning to be successfully introduced to medical devices (Corritore, Wiedenbeck, Kracher, & Marble, 2012). AI is expected to revolutionize the way medicine is practiced, improve process efficiency and thus change the way health care is provided (Kassam & Kassam, 2020).

The next area is the clinical decision-support system (CDSS), defined as “any electronic system designed to aid directly in clinical decision-making, in which characteristics of individual patients are used to generate patient-specific assessments or recommendations that are then presented to clinicians for consideration” (Kawamoto, Houlihan, Balas, & Lobach, 2005). CDSSs apply the best-known medical knowledge to patient data to generate case-specific decision support, particularly in preventive care services and treatment recommendations (Bright, et al. 2012). The technology allows cheap and effective diagnostics, detecting any deviations from standards, analyzing the results by comparing them with the huge amount of results stored in the database. The development of such systems is essential in countries like Poland, where only one in three visits to a primary care physician is justified by health problems. In many situations, the patient goes to the doctor only to be reassured that he or she is healthy (Grzela & Kuczyńska, 2019). CDSS solutions allows a doctor to quickly and effectively determine who needs medical consultations and additional diagnostic procedures under the supervision of a specialist. This can lead to savings in time and resources for doctors, patients, and medical facilities.

Aside from the many technical limitations of current AI technologies in comparison with human vision, language processing and context-specific reasoning, other distinctive challenges exist (Reddy et al., 2019). Despite the growing use of automated AI-based solutions by patients, social skepticism

is still noticeable regarding these applications (Hengstler, Enkel, & Duelli, 2016). On the one hand, according to research conducted in Great Britain, 63 percent of adults are concerned about the use of sensitive data, especially by artificial intelligence, to improve health care (Fenech, Strukelj, & Buson, 2018). However, in the situation of technological development and the disappearance of the human element in health care, one can speak of the emotional cost of this technological progress. A critical foundational relationship in health care interactions is the trust between patients and doctors (LaRosa & Danks, 2018). Patients are accustomed to interacting with people at various points of contact in the health care service and then taking into account their opinions, preferences and situations in decisions, interest in their emotional state, and providing support in this regard for health care services (Siemens Healthineers, 2019). Patients are skeptical about the new situation in which they need to communicate with a VA instead of an empathic and trusted person. According to Syneos Health Communications (2018) research, less than 20 percent of patients see any benefit to receiving a diagnosis or treatment recommendations from a virtual assistant. The black box nature of AI produces a shiver of discomfort for many people. How can we trust our health, let alone our very lives, to decisions whose pathways are unknown and impenetrable? (Feldman, Aldana, & Stein, 2019).

CURRENT RESEARCH

Hypothesis Development

Although AI systems are increasingly appearing in the area of health, people seem to be suspicious about such solutions. Their fears arise from their belief that the machine cannot recognize the uniqueness of the symptoms and features of the disease (Longoni, Bonezzi & Morewedge 2019). Patients are convinced that they are unique and different from everyone else, as individuals. Research also shows that relying on the diagnosis of a doctor (instead of a computer) reduces the patient's sense of responsibility for the decision (Promberger & Baron, 2006, study 1). In turn, receiving a recommendation from a computer increases the tendency to search for information (Promberger & Baron, 2006, study 2). A human decision provider is preferred, even when people are aware that computers can outperform humans (Dietvorst, Simmons, & Massey, 2015). It was also demonstrated that patients perceive those physicians who don't use computer-based solutions to provide a decision more positively (Shaffter, Probst, Merkle, Arkes, & Medow, 2012). Patient resistance to AI-based solutions may also be due to a digital environment that limits trust and contributes to people's greater suspicions (Taddeo, 2009; Ferrario, Loi, & Vigano, 2019).

In this paper, it is anticipated that using AI-based systems in health care evokes fears and concerns in patients, leading to lower willingness to rely on a medical diagnosis provided by a Virtual Assistant instead of a human. As a consequence, users exposed to a medical diagnosis provided by a Virtual Assistant would be more motivated to search for information and to confront a real doctor with that recommendation. In this paper we aim to investigate those assumptions regarding patients' fears with the use of non-declarative techniques related to eye movements (Henderson & Hollingworth, 1998). It is assumed that if participants are more suspicious about a diagnosis from VA, as their goal-orientation is focused on seeking more information, they will process information more deliberatively. This, in turn, can lead to different task-related attentional processes as measured by eye fixation patterns (Orquin & Loose, 2013). It is hypothesized that:

Hypothesis One: Participants presented with the diagnosis from a virtual assistant spend more time reading the diagnosis.

Hypothesis Two: Providing the diagnosis from a virtual assistant enhances focusing participant's attention on the signature of the diagnosis provider.

Hypothesis Three: Elements of the mobile app that are related to reducing concerns (contact button) gain more eye fixation in the virtual assistant condition.

In this study, the top down control of attention perspective is adopted, assuming that attention is driven by goal-orientation (Theeuwes, 2010). The second perspective - bottom up control of attention, emerging from the visual and textual properties of the stimuli - is rejected because participants saw the same stimuli and the only differentiating factor was the source of information. To justify our hypothesis, we rely on evidence that the uncertainty and the need to acquire new information can influence eye movement (Gottlieb, Hayhoe, Hikosaka & Rangel, 2014), and that fixation patterns reflect different cognitive goals and vary due to the variety of tasks or problems that need to be solved (Hayhoe, 2004).

In order to test the proposed hypotheses, we designed an eye-tracking experiment. The current study aims to measure attentional processes (reflected by eye movements) toward specific visual elements of the mobile app. To this purpose, we define areas-of-interest (AOIs) regarding the content of diagnosis, the signature of its source and contact-a-doctor button, imitating the possibility to contact with a real doctor.

RESEARCH OVERVIEW

In the current research, we invited people aged 18-29 from Poland to take part in the experiment. This group was chosen for at least two reasons. Patients are generally disappointed with the health care system in Poland, and people aged 18-29 assert the lowest level of satisfaction with the Polish health system. They mistake the waiting time for the service and the time spent consulting with doctors. They are also more sensitive than the other age groups to the intelligibility of the information provided by staff and the possibility of establishing a dialogue (Siemens Healthineers, 2019).

We presented them with a fictional scenario in which we exposed them to information about the mobile medical application and a fictional diagnosis from that application. We decided to use the existing health care application - StethoMe, which was introduced by a Polish start-up (Grzywalski et al., 2019). This application is an AI-based system that a patient can use to continuously monitor the state of the body. Patients are dealing with two manifestations of AI: an algorithm that analyzes and interprets the test results, and then communicates the result to the patient and manages his communication with the doctor. The application screens were modified for the purpose of this study - it did not conduct any medical measurement on the participants (they were only exposed to information about it and the results of an example medical measurement). Additionally, this application was virtually equipped with Evie VA, based on real VA offered by Evie company (Kot & Leszczyński, 2020).

We set two experimental conditions, differing in the source of the diagnosis provider. Half of the participants (graduate and undergraduate students of management) saw the diagnosis from a human (doctor) while the other half saw the one from the VA. Participants were randomly assigned to one of two conditions. The experiment lasted for about 10 minutes. It consisted of a stimuli exposition (the first screen was related to initial information about the mobile app and the second to medical examination and diagnosis) and a battery of questions related to personal characteristics and metrics. In order to measure participant's visual attention and cognitive effort, we used the screen-based eye tracking device, capturing gaze data at 120 Hz. A 9-point calibration was carried out before the experiment started. When the experiment finished, participants received information about the purpose of the study and were debriefed.

Experimental Design and Stimuli

In order to test the hypotheses proposed, we used a single-factor experiment confined to two groups. The independent variable was the source of the medical diagnosis (human vs. VA). When we exposed participants to the first screen, which contained an initial piece of information about the medical app, the manipulation included a short description of the main features of the app. The human-condition information contained claims that (1) a real doctor (we used a fictional name: "Ewa Gran") will carry out any assistance that is needed, (2) a real doctor is responsible for sending the diagnosis to

the user and (3) a real doctor will provide all information in the app. In the VA-condition, all the features remained identical, except for the claim that a VA (we used a name: “Evie”) is responsible for (1) providing any help, (2) sending the diagnosis, and (3) providing all the information. The first screen contained other information related to the necessary information about the application, and that remains consistent through all experimental conditions.

The second screen presented to the participants contained results of the health examination and information that one of these results slightly exceeds the norm. Regarding the results, participants saw diagnosis from a human (vs. VA), suggesting that exceeding the norm is not significant, and contact with a doctor is not necessary. We also informed participants that they might use the application to contact a human (vs. VA) to obtain more information. The only manipulation used in the second screen was related to the signature below the diagnosis and the information about obtaining additional information. Participants from the human-condition saw a signature from “Ewa Gran - the doctor” while those from VA-condition saw one from “Evie - virtual assistant”. The research procedure is illustrated in Figure 1.

Figure 1. Research procedure

Source: Author

Eye-Tracking Measurements

The data considered for the analysis was the time spent reading the information and the fixation count. Time spent (dwell time) offers information about the amount of time that a participant has spent looking at a particular area-of-interest (AOI). A longer duration of time spent on exploring AOI may indicate one’s motivation or the attention paid to a process stimulus (Farnsworth, 2018). In the current study, it is assumed that time spent on reading information about VA (vs. human) is longer, as participants are being exposed to a novel situation, and the duration of looking reflects their motivation to understand it. This assumption is supported by an extensive body of research regarding ease of comprehension reflected by time spent looking at an object and eye fixations (Just & Carpenter, 1980; Kliegl, Nuthmann, & Engbert, 2006; Rayner & Reingold, 2015). For example,

Becker (2011) demonstrated that both perceptual and conceptual difficulty in processing lengthen dwell time. This finding is supported by experiments conducted by Liu, Cole, Belkin, Gwizdka, and Zhang (2011) suggesting that longer dwell time indicates processing difficulty.

The fixation count is the second metric used in the current study to capture goal-directed eye movement (Krauzlis, Goffart, & Hafed, 2017). As emphasized in hypothesis 3, it is expected that VA participants are more willing to contact a real doctor, so they focus more on that button. The fixation count informs researchers about attention and interest in the button, and, as this metric is goal-directed, it is assumed that it reflects the behavioral intention to contact a real doctor (Susac, Bubic, Kaponja, Planinic, & Palmovic, 2014). For example, number of fixations is a well-used metric of the ease or difficulty during the performance of problem-solving tasks (e.g. Gegenfurtner, Lehtinen, & Säljö, 2011), allowing the prediction of behavior and preference on the basis on fixations (Orquin & Loose, 2013).

RESULTS

An Independent Two-Sample T-Test was conducted (due to the relatively small group sample, $N = 38$, 71% female; $M_{age} = 22$, $SD = 1.62$) to determine the statistical significance between the variables. The first step of analysis was regarding the perception of the introductory information about the app and its features. We took into consideration 3 AOIs related to essential claims (presented on the first screen) that differ across the experimental conditions (VA vs. human). Statistical analysis revealed no statistically significant differences between the VA and human groups regarding time spent on reading information and fixation count (Figure 2).

Figure 2. Mean time spent on reading app features

Source: Author

The second step of the analysis was crucial to the test of the hypotheses proposed. First, the assumption about the time spent on reading the diagnosis (H1) between two experimental conditions (VA vs. human) was tested. Contrary to the predictions, there is no statistically significant difference in time spent (in milliseconds) on reading due to the provider of the diagnosis, $M_{VA} = 5400.47$, $SD = 3490.10$ vs. $M_{human} = 6061.32$, $SD = 3205.21$, $t(36) = -.608$, $p = .401$. Second, the assumption about the attention paid to the signature of the diagnosis provider (H2) was tested. Again, contrary to the predictions, there is no statistically significant difference between participants exposed to diagnosis from VA vs. a human diagnosis as measured by fixation count, $M_{VA} = 3.21$, $SD = 2.44$ vs. $M_{human} = 3.27$, $SD = 2.22$, $t(36) = -.221$, $p = .836$.

The last analysis was conducted to test H3. It is expected that the contact-a-doctor button attracts more fixation among participants who are presented with the diagnosis provided by VA (vs. human). The results obtained confirmed this assumption. Fixation count is significantly greater in the VA condition, $M_{VA} = 4.37$, $SD = 3.89$ vs. $M_{human} = 1.79$, $SD = 2.64$; $t(36) = 2.392$, $p = .022$. Results are presented in Figure 4.

Figure 3. Mean time spent on reading diagnosis

Figure 4. Mean fixation count for 2 AOs related to: Contact-a-doctor button and Signature

** - statistically significant differences for $p < .05$
Source: Author

DISCUSSION

The results obtained demonstrate that the processing of introductory information about the medical app was the same across the two experimental groups. The analysis revealed that participants did not process information from the VA more eagerly than information from the real doctor (who provides the diagnosis to the user). It seems that at some initial stage of interaction with the mobile app, claims regarding the use of AI systems do not evoke any concerns from the participants of our study. This is contrary to other studies that suggests limited trust, high levels of initial resistance, and the perceived risk from radical innovations in the development process or at an early phase of diffusion (Heidenreich & Spieth, 2013). Moreover, with regard to the concept of e-trust (Taddeo, 2009), a digital environment does not meet the minimal requirements for trust to emerge. Thus, people exposed to AI-based systems might be more suspicious about relying on its recommendations (Ferrario et al., 2019). The dissimilarity of results in comparison to our research may be explained by the research sample, which consisted of relatively young participants who are more trusting in autonomous systems (Zhang, Na, Robert, & Yang, 2018). On the other hand, AI-based systems can be still novel for patients in the health care domain, so participants could interpret a virtual assistant as something atypical, thus causing concerns to arise. This assumption, however, has not been evidenced by the eye movement patterns.

The second step of data analysis covered visual processing of the screen related to examination and diagnosis. The 3 AOIs were defined according to the content of diagnosis, the signature of the diagnosis provider, and the contact-a-doctor button. Contrary to our assumption (hypothesis 1), participants presented with the diagnosis from a virtual assistant did not spend more time reading the diagnosis. Also, hypothesis 2 was rejected, as providing participants with the diagnosis from VA did not enhance focusing attention on the signature of the diagnosis provider. Crucially, a statistically significant difference in attention paid to the contact-a-doctor button was found during the fixation count. As expected (hypothesis 3), VA participants focused more on the button compared to those provided with the diagnosis from a real doctor. This is consistent with Promberger & Baron (2006) claiming recommendation from a computer enhance additional information search. In our scenario, participants could expect to receive more details from the doctor.

Despite the lack of differences in the processing of essential features of the health app and the diagnosis itself, VA participants focused more on the AOI related to communicating with a doctor. This result is surprising, because if interaction with VA evokes concerns about diagnosis reliability or accuracy, it should have led to more significant cognitive effort from the test subjects during processing. A potential explanation may relate to an implicit preference for personal contact with a doctor when a patient suspects his or her health might be at risk. Our participants were informed that their results slightly exceeded the norm. Thus, they might not be motivated to actively engage in the processing of the diagnosis. The discrepancy between the processing of diagnosis and focus on the contact-a-doctor button could result from a variation in perception of the risk to their health, confirming that attitude to health is driven by the one's locus of control (Stephoe & Wardle, 2001). Specifically, participants could deprecate the value of diagnosis information from both humans and VA, and not actively engage in processing it, as they felt that their health was not at risk. The difference in attention paid to the contact button could therefore result from a general preference for personal contact with a doctor evoked by the message from VA. This result might be explained by Longoni, Bonezzi and Morewedge (2019) findings that each patient perceives his or her diseases as unique, so they do not want to be treated by AI-based algorithm as just one of many similar cases. Therefore, they seek diagnosis from a doctor, who will treat them as individuals. This is also coherent with the research showing that the main fears patients have about AI are the lack of human oversight (Syneos Health Communication, 2018).

CONCLUSIONS

In the current study, the visual processing of the medical app providing participants with a fictional diagnosis of their heart activity was investigated. The results discussed above lead to three conclusions. First, performing a medical analysis by systems based on artificial intelligence is acceptable for these participants, unless the result of the examination is delivered by a real doctor (human). This result should be considered from the perspective of the skepticism associated with the high emotional cost of implementing automated solutions for patients. Patients place trust in their doctors and different medical professionals who must interact to support the care of a patient (LaRosa & Danks, 2018). The introduction of health care AI-based applications can potentially have significant impacts on those relationship, which are based on trust. Thus, we suppose that AI systems are unlikely to replace human clinicians on a large scale, but rather will augment clinician's efforts to care for patients, so they should be treated as assistive technologies. Over time, human clinicians may move toward tasks and job designs that draw on their uniquely human skills, such as empathy, persuasion and big-picture integration (Davenport & Kalakota, 2019).

Second, the results lead to questions about human-machine interface design and communication methods, which can affect the level of acceptance for autonomous technologies (Hengstler, Enkel, & Duelli, 2016) and reduce the emotional necessity of having to contact a human doctor. As the diagnosis in our study was presented without any room for human involvement, the acceptance of such technologies by patients goes beyond perceived usefulness, perceived ease of use and attitudes towards use (Song, 2019). The method and form of communication refers, to the issue of anthropomorphism, among other things, as described according to De Visser et al. (2016) adding human features (appearance or social ability) increases trust resilience. Trust to VA in health care can significantly affect the effectiveness of this type of solution.

Third, our observations may give an exciting premise for further research if extended to investigate behavioral intention and correlation with eye-tracking data. The behavioral intention would be reflected by eye-tracking data. Another issue is the aesthetic appraisal of the virtual assistant, which could differ due to the level of anthropomorphism (MacDorman & Chattopadhyay, 2016). This kind of research could contribute to a better understanding of the adaptation of AI in patient use of medical applications.

This research has some limitations that are typical for experimental research. Experimental designs create situations that are inherently unrealistic, and thus not crucial for participants, so the relative robust health of the participants in the study could affect their propensity to engage in the experimental scenario, and moderate their responses to the manipulations used. Therefore, this research is the initial study and requires replication. However, it is suggested that such applications may be a valuable in the case of the health care systems that have similar conditions to those of the Polish one. Because of the current situation of the lack of physicians, where many patients go to the doctor only to get assurance that he or she is healthy, diagnostic applications based on AI can contribute to reducing this particular problem.

ACKNOWLEDGMENT

Projekt finansowany w ramach programu Ministra Nauki i Szkolnictwa Wyższego pod nazwą Regionalna Inicjatywa Doskonałości w latach 2019 - 2022 nr projektu 004/RID/2018/19 kwota finansowania 3.000.000 zł.

The project financed within the Regional Initiative for Excellence programme of the Minister of Science and Higher Education of Poland, years 2019-2022, grant no. 004/RID/2018/19, financing 3,000,000 PLN.

REFERENCES

- Aborujilah, A., Mohammad, R., Al-Othmani, A. Z., Husen, M. N., Long, Z. A., Ali, N. A., & Murugan, V. A. (2020, January). *Adoption and continuance intention model of applying telemedicine technology in digital games addiction*. Paper presented at the 14th International Conference on Ubiquitous Information Management and Communication (IMCOM), Taichung, Taiwan. doi:10.1109/IMCOM48794.2020.9001732
- Becker, S. I. (2011). Determinants of dwell time in visual search: Similarity or perceptual difficulty? *PLoS One*, 6(3), e17740. doi:10.1371/journal.pone.0017740 PMID:21408139
- Bright, T. J., Wong, A., Dhurjati, R., Bristow, E., Bastian, L., Coeytaux, R. R., & Wing, L. (2012). Effect of clinical decision-support systems: A systematic review. *Annals of Internal Medicine*, 157(1), 29–43. doi:10.7326/0003-4819-157-1-201207030-00450 PMID:22751758
- Corritore, C. L., Wiedenbeck, S., Kracher, B., & Marble, R. P. (2012). Online trust and health information websites. *International Journal of Technology and Human Interaction*, 8(4), 92–115. doi:10.4018/jthi.2012100106
- Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6(2), 94–98. doi:10.7861/futurehosp.6-2-94 PMID:31363513
- De Keyser, A., Köcher, S., Alkire, L., Verbeeck, C., & Kandampully, J. (2019). Frontline service technology infusion: Conceptual archetypes and future research directions. *Journal of Service Management*, 30(1), 156–183. doi:10.1108/JOSM-03-2018-0082
- De Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology. Applied*, 22(3), 331–349. doi:10.1037/xap0000092 PMID:27505048
- Dietvorst, B., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology. General*, 144(1), 114–126. doi:10.1037/xge0000033 PMID:25401381
- Domagała, A., & Klich, J. (2018). Planning of Polish physician workforce: Systemic inconsistencies, challenges and possible ways forward. *Health Policy (Amsterdam)*, 122(2), 102–108. doi:10.1016/j.healthpol.2017.11.013 PMID:29208419
- Fanelli, G., Gall, J., Romsdorfer, H., Weise, T., & Van Gool, L. (2010). A 3-D audio-visual corpus of affective communication. *IEEE Transactions on Multimedia*, 12(6), 591–598. doi:10.1109/TMM.2010.2052239
- Farnsworth, B. (2018). *10 most used eye tracking metrics and terms*. Retrieved December 10, 2019, from <https://imotions.com/blog/7-terms-metrics-eye-tracking>
- Feldman, R. C., Aldana, E., & Stein, K. (2019). Artificial intelligence in the healthcare space: How we can trust what we cannot know. *Stanford Law & Policy Review*, 30, 399–419.
- Fenech, M., Strukelj, N., & Buston, O. (2018). *Ethical, social and political challenges of artificial intelligence in health*. Retrieved December 5, 2019, from <https://wellcome.ac.uk/sites/default/files/ai-in-health-ethical-social-political-challenges.pdf>
- Ferrario, A., Loi, M., & Viganò, E. (2019). *AI we trust incrementally: A multi-layer model of trust to analyze human-artificial intelligence interactions*. Philosophy & Technology.
- Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots: Concepts, design, and applications. *Robotics and Autonomous Systems*, 42(3-4), 1431–1466.
- Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, 23(4), 523–552. doi:10.1007/s10648-011-9174-7
- Gottlieb, J., Hayhoe, M., Hikosaka, O., & Rangel, A. (2014). Attention, reward, and information seeking. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 34(46), 15497–15504. doi:10.1523/JNEUROSCI.3270-14.2014 PMID:25392517

- Grzela, E., & Kuczyńska, E. (2019). Sztuczna inteligencja to rewolucja w medycynie- Relacja z. *Impact (Sydney)*, 19. Retrieved December 4, 2019, from <https://pulsmedycyny.pl/sztuczna-inteligencja-to-rewolucja-w-medycynie-relacja-z-impact-19-962162>
- Grzywalski, T., Belluzzo, R., Piecuch, M., Szajek, M., Bręborowicz, A., Pastusiak, A., & Kociński, J. et al. (2019). Fully interactive lungs auscultation with AI enabled digital stethoscope. In *Proceedings of Conference on Artificial Intelligence in Medicine in Europe*. Cham, Switzerland: Springer International Publishing AG. doi:10.1007/978-3-030-21642-9_5
- Hayhoe, M. M. (2004). Advances in relating eye movements and cognition. *Infancy*, 6(2), 267–274. doi:10.1207/s15327078in0602_7 PMID:33430534
- Heidenreich, S., & Spieth, P. (2013). Why innovations fail? The case of passive and active innovation resistance. *International Journal of Innovation Management*, 17(5), 1–42. doi:10.1142/S1363919613500217
- Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 269–293). Elsevier Science Ltd. doi:10.1016/B978-008043361-5/50013-4
- Hengstler, M., Enkel, E., & Duelli, S. (2016). Applied artificial intelligence and trust. The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change*, 105, 105–120. doi:10.1016/j.techfore.2015.12.014
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354. doi:10.1037/0033-295X.87.4.329 PMID:7413885
- Kashyap, R. (2020). Applications of wireless sensor networks in healthcare. In R. Kashyap (Ed.), *IoT and WSN applications for modern agricultural advancements: Emerging research and opportunities* (pp. 8–40). IGI Global. doi:10.4018/978-1-5225-9004-0.ch002
- Kassam, A., & Kassam, N. (2020). Artificial intelligence in healthcare: A Canadian context. *Healthcare Management Forum*, 33(1), 5–9. doi:10.1177/0840470419874356 PMID:31505963
- Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *BMJ (Clinical Research Ed.)*, 330(7494), 765–774. doi:10.1136/bmj.38398.500764.8F PMID:15767266
- Khan, R., & Das, A. (2018). Introduction to chatbots. In R. Khan & A. Das (Eds.), *Build better chatbots: A complete guide to getting started with chatbots* (pp. 1–11). Apress. doi:10.1007/978-1-4842-3111-1_1
- Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology. General*, 135(1), 12–35. doi:10.1037/0096-3445.135.1.12 PMID:16478314
- Kot, M., & Leszczyński, G. (2020). The concept of intelligent agent in business interactions: Is virtual assistant an actor or a boundary object? *Journal of Business and Industrial Marketing*, 35(7), 1155–1164. doi:10.1108/JBIM-10-2018-0291
- Krauzlis, R. J., Goffart, L., & Hafed, Z. M. (2017). Neuronal control of fixation and fixational eye movements. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1718), 1–12.
- LaRosa, E., & Danks, D. (2018). Impacts on trust of healthcare AI. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. New York, NY: Association for Computing Machinery. doi:10.1145/3278721.3278771
- Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies*, 77, 23–27. doi:10.1016/j.ijhcs.2015.01.001
- Liu, C., Cole, M., Belkin, N., Gwizdka, J., & Zhang, X. (2011, May). *Exploring the effect of task difficulty and domain knowledge on dwell times*. Paper presented at the Fifth Workshop on Human–Computer Interaction and Information Retrieval, Vancouver, Canada.
- Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *The Journal of Consumer Research*, 46(4), 629–650. doi:10.1093/jcr/ucz013

- MacDorman, K. F., & Chattopadhyay, D. (2016). Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition*, 146, 190–205. doi:10.1016/j.cognition.2015.09.019 PMID:26435049
- Maglio, P. P., & Spohrer, J. (2008). Fundamentals of service science. *Journal of the Academy of Marketing Science*, 36(1), 18–20. doi:10.1007/s11747-007-0058-9
- Mak, K. K., & Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. *Drug Discovery Today*, 24(3), 773–780. doi:10.1016/j.drudis.2018.11.014 PMID:30472429
- Marinova, D., De Ruyter, K., Huang, M. H., Meuter, M. L., & Challagalla, G. (2017). Getting smart: Learning from technology-empowered frontline interactions. *Journal of Service Research*, 20(1), 29–42. doi:10.1177/1094670516679273
- McDougall, R. J. (2019). Computer knows best? The need for value-flexibility in medical AI. *Journal of Medical Ethics*, 45(3), 156–160. doi:10.1136/medethics-2018-105118 PMID:30467198
- Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, 72, 432–440. doi:10.1016/j.chb.2017.02.067
- Nasirian, F., Ahmadian, M., & Lee, O. K. D. (2017, August). *AI-based voice assistant systems: Evaluating from the interaction and trust perspectives*. Paper presented at the Twenty-third Americas Conference on Information Systems, Boston, MA.
- Orquin, J. L., & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144(1), 190–206. doi:10.1016/j.actpsy.2013.06.003 PMID:23845447
- Promberger, M., & Baron, J. (2006). Do patients trust computers? *Journal of Behavioral Decision Making*, 19(5), 455–468. doi:10.1002/bdm.542
- Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P., & Mazurek, G. (2019). In bot we trust: A new methodology of chatbot performance measures. *Business Horizons*, 62(6), 785–797. doi:10.1016/j.bushor.2019.08.005
- Rajabion, L., Shaltooli, A. A., Taghikhah, M., Ghasemi, A., & Badfar, A. (2019). Healthcare big data processing mechanisms: The role of cloud computing. *International Journal of Information Management*, 49, 271–289. doi:10.1016/j.ijinfomgt.2019.05.017
- Rayner, K., & Reingold, E. M. (2015). Evidence for direct cognitive control of fixation durations during reading. *Current Opinion in Behavioral Sciences*, 1, 107–112. doi:10.1016/j.cobeha.2014.10.008
- Reddy, S., Fox, J., & Purohit, M. P. (2019). Artificial intelligence-enabled healthcare delivery. *Journal of the Royal Society of Medicine*, 112(1), 22–28. doi:10.1177/0141076818815510 PMID:30507284
- Russell, S. J., & Norvig, P. (2009). *Artificial intelligence: A modern approach*. Prentice Hall.
- Scherer, M. U. (2016). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology*, 29(2), 353–399.
- Shaffer, V. A., Probst, C. A., Merkle, E. C., Arkes, H. R., & Medow, M. A. (2013). Why do patients derogate physicians who use a computer-based Diagnostic support system? *Medical Decision Making*, 33(1), 108–118. doi:10.1177/0272989X12453501
- Shah, R., & Chircu, A. (2018). IOT and ai in healthcare: A systematic literature review. *Issues in Information Systems*, 19(3), 33–41.
- Siemens Healthineers. (2019). Patient experience in Poland. Retrieved December, 4 2019, from <http://www.siemens-healthineers.pl/raport>
- Song, Y. W. (2019). *User acceptance of an artificial intelligence (AI) virtual assistant: An extension of the technology acceptance model* (Unpublished doctoral dissertation). The University of Texas, Austin, TX.
- Stepoe, A., & Wardle, J. (2001). Locus of control and health behaviour revisited: A multivariate analysis of young adults from 18 countries. *British Journal of Psychology*, 92(4), 659–672. doi:10.1348/000712601162400 PMID:11762867

Susac, A. N. A., Bubic, A., Kaponja, J., Planinic, M., & Palmovic, M. (2014). Eye movements reveal students' strategies in simple equation solving. *International Journal of Science and Mathematics Education*, 12(3), 555–577. doi:10.1007/s10763-014-9514-4

Syneos Health Communications. (2018). *Artificial Intelligence for Authentic Engagement. Patient perspectives on healthcare's evolving AI conversation*. Retrieved March, 28 2020, from <https://syneoshealthcommunications.com/perspectives/artificial-intelligence-for-authentic-engagement>

Taddeo, M. (2009). Defining trust and e-trust: From old theories to new problems. *International Journal of Technology and Human Interaction*, 5(2), 23–35. doi:10.4018/jthi.2009040102

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, 135(2), 77–99. doi:10.1016/j.actpsy.2010.02.006 PMID:20507828

Weizenbaum, J., & McCarthy, J. (1976). *Computer power and human reason: From judgment to calculation*. W.H. Freeman and Company.

Wijesinghe, I., Gamage, C., Perera, I., & Chitraranjan, C. (2019, July). *A smart telemedicine system with deep learning to manage diabetic retinopathy and foot ulcers*. Paper presented at the Engineering Research Conference (MERCon), Moratuwa, Sri Lanka. doi:10.1109/MERCon.2019.8818682

Zhang, Q., Na, D., Robert, L. P., & Yang, X. J. (2018, May). *Trust in AVs: The Impact of expectations and individual differences*. Paper presented at the conference on Autonomous Vehicles in Society: Building a Research Agenda, East Lansing, MI.