

Are you aware of what you are watching? Role of machine heuristic in online content recommendations

Soyoung Oh, Eunil Park*

Department of Applied Artificial Intelligence, Sungkyunkwan University
Seoul, Republic of Korea
sori424@g.skku.edu, eunilpark@skku.edu

Abstract

Since recommendation systems have been created and developed to automate the recommendation process, users can easily consume their desired video content on online platforms. In this line, several content recommendation algorithms are introduced and employed to allow users to encounter content of their interests, effectively. However, the recommendation systems sometimes regrettably recommend inappropriate content, including misinformation or fake news. To make it worse, people would unreservedly accept such content due to their cognitive heuristic, machine heuristic, which is the rule of thumb that machines are more accurate and trustworthy than humans. In this study, we designed and conducted a web-based experiment where the participants are invoked machine heuristic by experiencing the whole process of a machine or human recommendation system. The results demonstrate that participants ($N = 89$) showed a more positive attitude toward a machine recommender than a human recommender, even if the recommended videos contain inappropriate content. While participants who have a high level of trust in machines exhibited a negative attitude toward recommendations. Based on these results, we suggest that a phenomenon known as algorithm aversion might be simultaneously considered with the machine heuristic in investigating human interaction with a machine.

Introduction

As online content recommendation systems have penetrated deeply into the daily use of online technology, service providers have made significant efforts to provide appropriate content that meets users' preferences. One solution involves employing a unique and robust algorithm to address the user's needs via system-initiated personalization. For instance, YouTube, the world's most popular online video-sharing platform, provides the auto-play button, which plays another video once the current video is finished by predicting what the users would like to see based on their viewing history (Bedjaoui, Elouali, and Benslimane 2018). However, the convenience of such automation makes it inevitable for users to face several problems (Pariser 2011; Tufekci 2018). One of the notable issues is the "filter bubble", which results in biased recommendations on specific political opin-

ions (Nie, Zhang, and Liu 2014). Owing to their customization, users would be exposed to the content they favor resulting in narrowing users' chances of exploring other opinions (Pariser 2011). For example, in online media channels, several news algorithms overrepresent right-wing voices, which can be characterized by hateful vocabulary, violent content, and discriminatory biases, while under-represents others (Otoni et al. 2018). Moreover, the auto-play button, which restricts users to perform any additional activities to watch the next video, can create an infinite loophole of presenting more biased videos (Tufekci 2018). For instance, videos about vegetarianism led to videos about veganism, and videos about jogging lead to videos about running ultramarathons. Such a problem is related to the revenue-making system of these platforms. As higher views produce higher revenue, it is inevitable for the platforms to recommend content that users would like to increase viewership in the platforms (Fleder and Hosanagar 2009; Hosanagar et al. 2014).

However, such extreme recommendations can prohibit individuals from being adequately informed about the status quo and rational thinking (Haim, Graefe, and Brosius 2018). Moreover, these problems can be accelerated with algorithmic cues of the recommendation systems. As humans are known as cognitive misers who tend to make decisions on the basis of cognitive heuristics, which can be triggered by contextual cues in the situation or on the interface (Sundar 2008). Along with the machine-like cues in the recommendation systems, the machine heuristic, the rule of thumb that machines are more objective and trustworthy than humans (Sundar and Kim 2019), can contribute to creating a positive attitude toward the algorithmic recommendations. That is, the interface cues trigger the humans' instinctive preference for machine agents over human agents to lead them unconsciously consume the recommended videos without noticing inappropriate contents within the recommendations. In this paper, we aim to test this probability by answering the research question (RQ):

- **RQ:** Are there notable differences in users' attitudes toward machine-recommended and human-recommended videos, which contain inappropriate contents?

Moreover, considering several prior research on the notable roles of source credibility (Stroud and Lee 2013; Wölker and Powell 2018), we assumed that credibility

*Corresponding author.

which is led by recommendation cue would positively affect the attitude toward recommended videos. Thus, the following hypothesis is proposed:

- **H1:** When people find the recommendation system more credible, depending on whether the video is machine-recommended or human-recommended, they have a more positive attitude towards recommendations.

Trust is a critical factor that impacts not only face-to-face interpersonal relationship, but also important for human-machine communications (Yu et al. 2018; Hoff and Bashir 2015; Kim, Park, and Sundar 2013; Lee and Moray 1994). Particularly for the systems where users are required to make decisions based, at least partially, on machine recommendations. The experiential effects based on the context-dependent variables in machine system can directly influence the trust (Cohen et al. 1997). For example, the type of system, complexity, and the difficulty of the task determines the trust in automated system. In addition to the external factors, as humans are creatures of experience, they can use pre-existing knowledge from past interactions with automation when assessing the trustworthiness of novel systems (Hoff and Bashir 2015). The initially learned trust built on prior experience with the system or similar technology (Yuviler-Gavish and Gopher 2011) and preexisting attitudes, as well as expectations (Abe and Richardson 2006), can alter the formation procedures of the trust and subsequent decisions during interacting with automated systems. As machine heuristic is primed by the cognitive accessibility at the time of decision making (Higgins, Bargh, and Lombardi 1985), the users' predispositions of the trust in the source of the recommendation would interact with the recommendation cues. In other words, the effect of reliance on a machine-like recommendation cue is dependent upon prior experiences with machines. However, this effect will not be found when the human recommendation cue is presented. Thus, we propose the following hypothesis:

- **H2:** The higher level of predispositions of trust in machines allows users to have more positive attitude to machine-recommended videos, which is not the case in human-recommended videos.

Based on the research question and hypotheses, the research model is developed to illustrate the role of machine heuristic in credibility judgment and attitude (Figure 1).

Related Work

The concept of machine heuristic roots back in the psychological underpinnings of automation bias. The automation bias describes errors made when decision-makers rely on automated cues as a heuristic replacement for information seeking and processing (Mosier et al. 1996). Potential negative effects of this cognitive bias can be separated into automation commission errors (i.e., overreliance on flawed systems or failure to detect a misleading signal), and automation omission errors (i.e., failure to act on an accurate signal). Several scholars have found notable evidences of these errors in various domains, including aviation (Layton, Smith, and McCoy 1994; Mosier et al. 1996; Skitka,

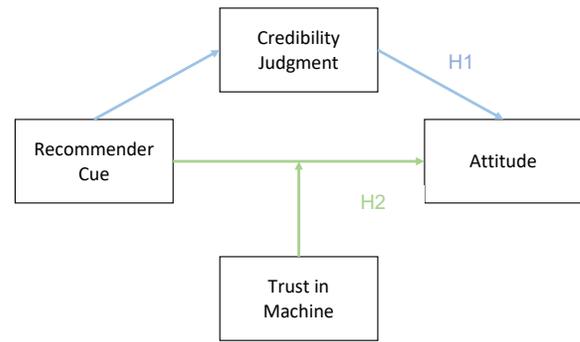


Figure 1: A proposed research model to analyze machine heuristic in recommendation system

Mosier, and Burdick 1999), life-support systems (Wickens et al. 2015), and healthcare (Wolfstadt et al. 2008; Bogun et al. 2004; Goddard, Roudsari, and Wyatt 2012). Likewise, under high task load conditions, automation-induced complacency, a self-satisfaction that may result in non-vigilance based on an unjustified assumption of a satisfactory system state, may occur (Parasuraman, Molloy, and Singh 1993). This psychological state can lead to danger in both experts and novice operators of automated systems, particularly with highly reliable systems (Parasuraman, Molloy, and Singh 1993; Singh, Molloy, and Parasuraman 1997; Bagheri, Jamieson et al. 2004). Because user experience and practice do not appear to mitigate automation complacency, it is possible that the complacency may be driven by initial high trust in the automation (Parasuraman and Manzey 2010).

Based on such ingrained faith in automated operations, the machine heuristic may result in attributions of randomness, objectivity, and other mechanical characteristics to the machine-like interface (Sundar 2008). Despite the extensive use of automation in computing systems, little research has been explicitly conducted to provide evidence for the operation of the machine heuristic in domains where ethical use of automation is required. In the context of online transactions involving a sense of security, Sundar and Kim (2019) found out that the users tend to disclose their personal financial information to automated decision aids rather than human beings. Considering the sensitive nature of human behavior, especially the study of human cognition, it is important to bridge the gap between the experiment and the real world. However, previous study with machine heuristic assessed users' behavioral intention in the form of self-reports, after they were asked to imagine that they were in the scenario portrayed in the screenshot (Sundar and Kim 2019). That being said, there might be a discrepancy between their behavioral intention and actual behavior when it comes to reality. Therefore, to fill this gap with empirical evidence, we devised an experimental stimuli that can capture the behavioral data of users in experiencing the automated system instead of asking them to report their intention.

Experiment

Method

We received the institutional review board (IRB) approval from the affiliated university (2020-11-025). A between-subjects experiment was conducted to validate the hypotheses and the research model. Participants experienced to one of the recommendation systems (i.e., human versus machine). They were then instructed to respond to a survey, measuring users' degree of trust in the machine, credibility of the system, and attitude toward the recommended video list, which can be organized by appropriate and inappropriate contents. Our experimental stimulus can be replicated by using publicly available codes¹.

Participants

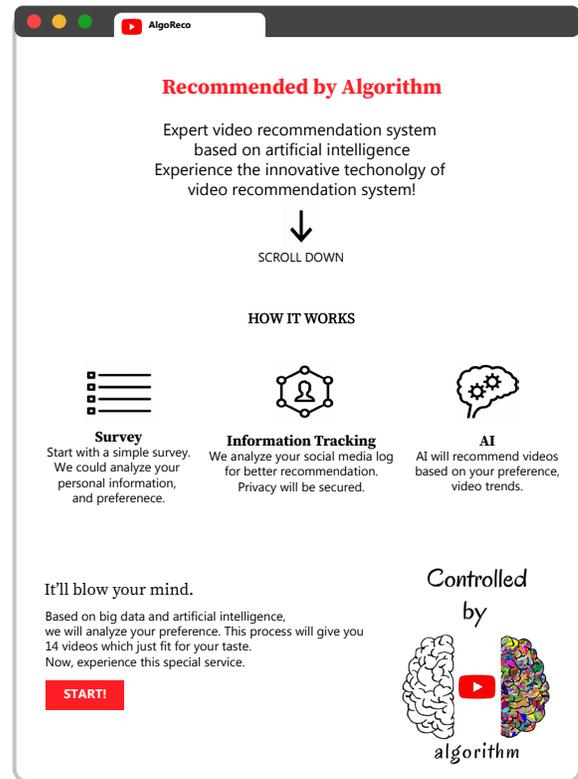
A total of 90 participants were recruited from one of the South Korean professional survey agencies. After a filtering procedure, 89 validated responses were organized by 35 males (39.3%) and 54 females (60.7%). The age ranged from 20 to 60 years. While 43.8% were in their 20s, 32.6%, 15.7% and 7.9% were in 30s, 40s, and 50s, respectively. The educational background ranged from below high school graduate to Ph.D., and the majority of the participants had a bachelor's degree (60.7%). Half of the participants were assigned to the machine recommender condition and the others were assigned to the human recommender condition.

Experimental Stimulus and Procedure

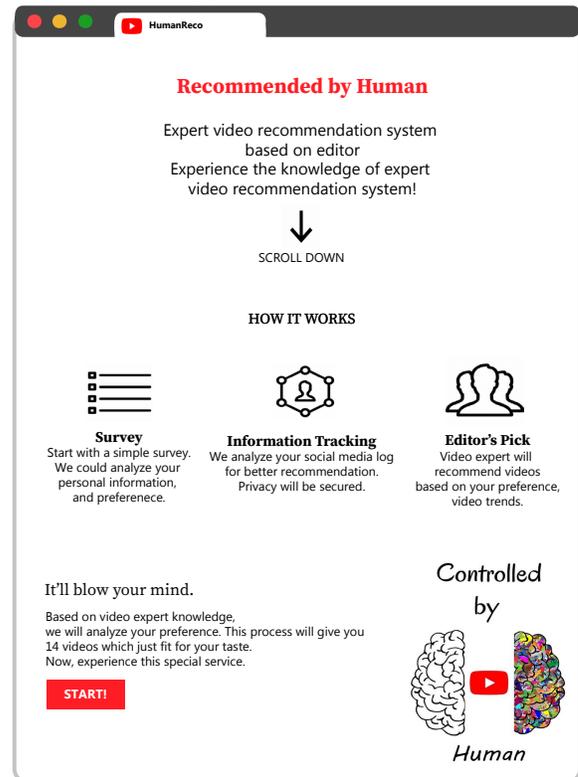
We designed a web-based recommendation system by adapting YouTube interface as in Figure 2. Different recommender identity cues were employed for human and machine conditions (Koh and Sundar 2010; Kim 2016).

For the human condition, we proposed the recommendation is executed by an expert video content editor. While we described that artificial intelligence recommends the videos in the machine condition. Details of the whole procedure of the recommendation in each condition are as follows. On the introductory page, participants in the human condition were provided an explanation that human editors recommended the videos, while those in the machine condition were informed that the videos were recommended by machine editors. Participants in the human condition were shown a human editor's image on the recommendation loading page and saw the message that the human editor took part in the recommendation procedures. In the machine condition, participants saw a machine editor's image and received the notification that the algorithm provided the recommendations. Moreover, pre-recorded voices were presented as cues for recommender identities, such that participants in the human condition were exposed to a human voice saying "a video specialist is in the process of generating a list of videos for you". Participants in the machine condition were exposed to a synthetic voice saying "an algorithm specialized in video recommendations is in the process of generating a list of videos for the participant". Female voices, which were neutral and monotone in style, were used in both conditions.

¹<https://github.com/sori424/Machine-Heuristic-Experimental-Stimulus>



(a) Machine condition



(b) Human condition

Figure 2: Introduction web-pages for each condition

Participants were then instructed to answer questions regarding demographic information which have no effect on the recommendation performance. This step's purpose was to increase credibility of the recommendation system by inducing each participant to believe that the system uses demographic data to analyze user preferences and provide personalized recommendations. An identical list of recommended videos was given to both human and machine conditions, and seven out of the 14 videos in the recommended list contained inappropriate content. The videos' inappropriateness was assessed by the video violation guideline in YouTube². The inappropriate videos included content related to assault, hate speech, fake news, pranks with a threat, and how-to guides of illegal activities (e.g., how to unlock a bicycle chain and how to make guns with a 3D printer). A manipulation check question ("Please list down the inappropriate videos you remember") was employed to check the inappropriateness of the videos during the pilot test with 15 participants. Six of the fifteen respondents indicated that they saw severely inappropriate videos in the recommended list. Finally, the participants were debriefed about the study's purpose upon completing the experiment.

The entire procedure of the experiment was as follows:

- (i) Each participant was instructed to read all descriptions of the recommender on the introductory web page and click the start button at the bottom of the page.
- (ii) A survey on users' demographic information was given: name, e-mail address, nationality, age, gender, video interest, and social media user ID were asked.
- (iii) The participant was advised to wait for the video recommendations as they continued watching the loading page. The loading page contained visual and auditory cues for triggering users to be aware of the recommender identity.
- (iv) An identical list of pre-selected videos was suggested to the participants. Depending on the conditions, either an image of a machine or a human was placed on the top of the page.
- (v) The participant was given two minutes to click and watch the recommended videos on their own choices.
- (vi) The participant received the confirmation codes to verify that they followed the whole recommendation process to participate in the original survey.
- (vii) Upon completion of the survey, the participants were debriefed about the study's purpose. They were made aware that the video list presented in the experiment was not recommended by the real system and the information that they inputted was not saved.

Measures

All questionnaire items were measured using a 7-point Likert scale (1 = strongly disagree; 7 = strongly agree).

Credibility Credibility was measured using four items: 1) The recommendation system has expertise in recommending videos, 2) The recommendation system knows well about

²<https://creatoracademy.youtube.com/page/course/community-guidelines>

the video contents, 3) The recommendation is reliably qualified, and 4) The recommendation system is inexperienced (reversed). The Cronbach's alpha for all four items, which were validated by prior online service research (Featherman, Miyazaki, and Spratt 2010), was 0.89.

Trust in machine Trust in machine was measured using three items: 1) Machines are more objective than human when performing a task, 2) Machines make better decisions than humans when it comes to ethics, 3) The decision of a machine does not cause damage to others. The Cronbach's alpha for three items, which were validated by prior machine heuristics research (Sundar and Kim 2019), was 0.56.

Dependent variable Attitude toward the recommendation system, was assessed by four items: 1) I am satisfied with the recommendation system, 2) Using this recommendation system is useful, 3) Using this recommendation system is a good idea, 4) It would be uncomfortable if I cannot use this recommendation system. The Cronbach's alpha for these items, which were validated by prior public perception research (Kim 2016), was 0.92.

Results

Manipulation check

To determine the effectiveness of manipulation in an experimental stimulus, we used different questionnaires in each condition as follows. For the machine condition, participants were instructed to respond to the following agreement statement by 7-point Likert scale: 1) "I had a video recommendation from a machine editor" ($M = 4.73$, $SD = 1.38$).

For the human condition, participants responded to the following statement: 2) "I had a video recommendation from a human editor" ($M = 4.16$, $SD = 1.29$). The participants in machine condition were found to be more likely to recognize the source of the recommendation compared to the human condition ($t(87) = 2.00$, $p = .047$).

To explore the participants' engagement by using a 7-point Likert scale in the experiment (Koh and Sundar 2010), they were instructed to respond to the following items: 1) "How important was the recommendation task for you?" 2) "How involved were you in the experiment?" 3) "Did you feel as though you were experiencing an actual recommendation system?". No significant difference between machine ($M = 3.89$, $SD = 1.18$) and human conditions ($M = 3.48$, $SD = 1.37$) was observed. Participants in both conditions were found to have been equally motivated to participate in the experiment ($t(87) = 1.49$, $p = .14$).

Hypothesis Tests

An independent t -test was conducted to address the research question. The result showed the significant difference across attitude toward recommendation system ($t(87) = 1.99$, $p < .05$). Specifically, the participants in the machine condition ($M = 4.18$, $SD = 1.35$) were more likely to have a positive attitude toward the recommended videos than those in the human condition ($M = 3.62$, $SD = 1.26$). Besides, the machine cue triggers the participants' positive attitudes toward recommendations ($\beta = 1.06$, $p < .05$), but the human cue

which indicates the video expert does not affect to the participants' attitudes toward the recommendations ($p = .86$).

Then, we conducted mediation and moderation analysis by using Hayes's PROCESS macro (Model 5) with 5,000 bootstrap samples to answer the hypotheses that we suggested (Hayes 2017). The model was bias-corrected and set at a 95% confidence interval to analyze each machine's proposed research model and the human condition. All control variables were considered as co-variants. The results are summarized in Figure 3.

To validate the H1, which addresses the mediating role of credibility, a simple regression analysis was conducted. The results show that there is a direct effect of machine cue on credibility ($SE = 0.11$, $\beta = 0.40$, $p < .01$, $CI = [0.16, 0.64]$). That is, machine cue triggers participants to perceive the recommendation system as credible. Besides, the increased credibility dependent on the machine cue leads to have positive attitude toward recommendations ($SE = 0.11$, $\beta = .29$, $p < .01$, $CI = [0.09, 0.53]$). The human cue also had a direct effect on credibility ($SE = 0.35$, $\beta = 0.35$, $p < .01$, $CI = [0.10, 0.60]$), where human cue makes the recommendation system more credible. Moreover, as machine condition, increased credibility contributes to positive attitude toward human recommender's recommendations ($SE = 0.11$, $\beta = .18$, $p < .01$, $CI = [0.01, 0.46]$). Regardless of the interface cue, the higher the credibility, the greater the positive attitude toward recommendations. H2 hypothesized that a positive attitude toward recommendations will be presented among the participants with stronger trust in machines in the presence of machine cues in the interface, rather than a human agent. There is a significant effect of trust in machine on attitude in presence of machine cue ($SE = 0.10$, $\beta = -.27$, $p < .05$, $CI = [-0.48, -0.06]$), while no significant effect of the trust in machine on the relationship between human cue and attitude is presented ($p = .59$).

Meanwhile, the effect of trust in machines on the relationship between machine cues and attitude is presented in an opposite direction from our expectation. That is, the participant with stronger trust in a machine, who saw machine cue, has a negative attitude toward recommendations. Therefore, we conducted further analysis to investigate the conditional effects of the variable. The results showed that as trust in the machine and perception that the machine as a main source of the recommendation increased, the attitude becomes more negative ($SE = 0.10$, $\beta = -0.35$, $p < .01$, $CI = [-0.57, -0.13]$). However, there was no significant relationship between attitude and low ($SE = 0.14$, $\beta = 0.11$, $p = .44$, $CI = [-0.18, 0.40]$) or average ($\beta = -0.12$, $p = .19$, $CI = [-0.30, 0.06]$) trust in machine. The interaction points are plotted in Figure 4. Participants with low trust in machines tend to have a positive attitude toward the recommended videos when they are more likely to perceive the recommendation system as a machine. For participants with a more than average level of trust in a machine, the increased level of perceiving the recommendation system as a machine led to a negative attitude toward the recommended videos. That is, the negative attitude can be explained by the participants' dissatisfaction with the recommendations. The participants who have high trust in the machine would expect the performance of

the machine recommender system to accurately predict their video preferences. However, as the recommendation list was created regardless of the preferences of the participants, the dissatisfaction might negatively affect the attitude.

Conclusion

This study contributes to providing an overview of users' actual experiences when facing a similar situation by conducting the experimental study. As one of the empirical pieces of evidence in addressing cognitive heuristic approaches to online recommendation systems, this study presents the effects of machine-related interface cues on users. In specific, our result from the web-based experiment confirmed that people are inclined to exhibit a more positive attitude toward machine-generated recommendations than human-generated recommendations. Even the recommendation video list in our experiment contains videos that violate the content guideline, the participants show cognitive heuristics in terms of revealing trust and reliance in automated suggestions. Besides, the higher the users perceive the recommendation as a machine, the users are more likely to have positive attitudes. From these results, we confirmed that the machine heuristic is involved in building a positive attitude to the recommended videos by algorithms. That is, merely identifying the artificial intelligence as the informant of the interactions can trigger the heuristic based on stereotypes about the operations of machines (Sundar 2020). Such automation bias induces people to adhere to automated decision authorities even if the outputs seem to violate the common sense of what is right (Mosier et al. 1996). With automated cues, people would not verify the given recommendations due to over-credence on the recommendation algorithm. To make matters worse, our results show that regardless of whether an agent is human or automated, the higher credibility of the online system leads to a more positive attitude toward the system. That being said, an authority heuristic could occur by using a well-known recommendation system's interface, YouTube (Sundar 2008). As the web interface generated for the experiment in this study reveals itself as having an official authority, the participants may confer believability to the recommended content with high credibility to a recommendation system.

Therefore, the participants would unconsciously consume inappropriate content due to the combined effect of both heuristics. That is, machine cue and its credibility assigned from authority could be the factors that contribute toward consuming misleading and harmful content by making people believe the recommendation performance without in-depth consideration. In this regard, users always should be aware of the recommended contents rather than depending on the various interface cues including algorithmic agents and their credibility.

While contrary to the findings of prior research, which suggested that the users' greater trait trust in machines leads to a higher positive attitude toward the performance of the machine (Sundar and Kim 2019), the results of the experiment showed contrasting results. The greater level of users' trust in machines leads to the more negative attitude toward the recommendations in the presence of machine cues. This

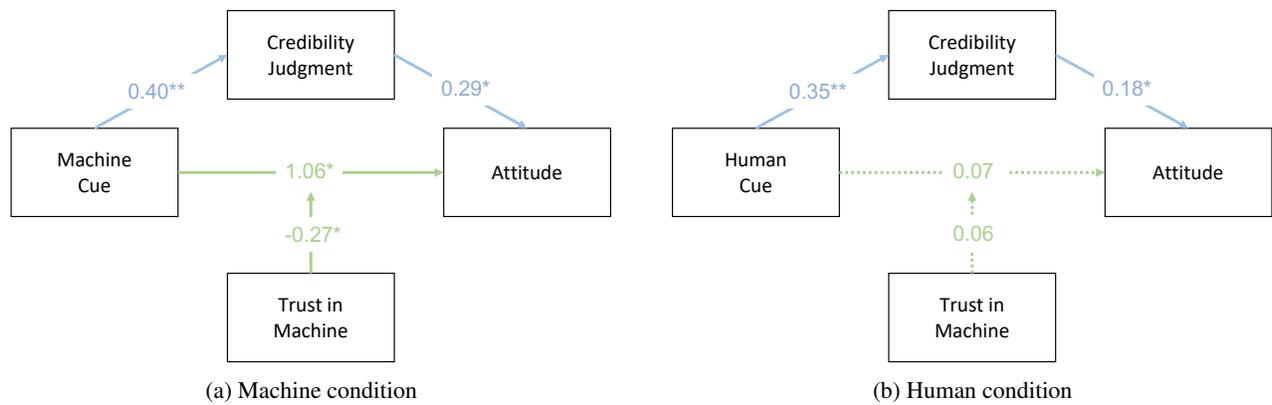


Figure 3: Results of the research models from each condition: machine and human recommender systems ($*p < 0.05$; $**p < 0.01$)

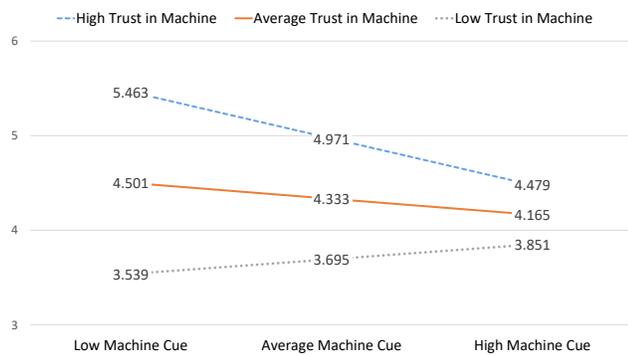


Figure 4: An interaction plot; each line indicates trust level in a machine, x-axis as a degree of perceiving the recommendation system as a machine, and y-axis as an attitude toward algorithmic recommendation system.

can be explained by users' general experiences of algorithmic operations; machine cues may have set high expectations on recommendations for users who have high trust in machines (Mosier et al. 1996). However, when a user realizes that a specific algorithm fails in a specific task, the cueing machine heuristic can result in negative reactions, such as algorithm aversion (Dietvorst, Simmons, and Massey 2015; Dzindolet et al. 2002). Dietvorst, Simmons, and Massey (2015) found that the errors that we tolerate in humans become less tolerable when machines make them where people have a general tendency to quickly lose confidence in algorithms after seeing them make a mistake. Hence, the participants with a firmer trust in machines would have expected the algorithmic recommendation system to predict the videos they were interested in with some degree of accuracy. Since the video list had been made in advance, the interests of the participants were not considered. Therefore, after receiving the recommendation, the participants might have experienced negative dis-confirmation between their expectation of the machine's performance and experienced performance, which led to negative attitudes.

Based on the result, we emphasize the necessity of expanding a research model to explore interactions between humans and machines in a wide range of forecasting domains. In addition to cognitive heuristics to explain humans' preferences in using algorithms in prediction tasks, other factors such as algorithm aversion can be considered to explore why people are dissatisfied with the predictions at the same time (Dietvorst, Simmons, and Massey 2015). In other words, multiple factors that trigger interactions in a different direction can be considered simultaneously.

Limitations

Although some academic findings have been presented, several limitations should be addressed in future studies. First, the potential biases regarding the recommendation system might affect the attitude of participants. As human-based video recommendation systems are not usual, users' perceived familiarity with machine recommenders may have positively affected attitudes. This can be assumed by the manipulation check result where the participants in human condition were less likely to recognize the source of recommendation compared to the machine condition. We mitigated this issue by employing the same well-known video recommendation interface for both conditions, familiarity coming from the recommendation algorithm could have positively affected the machine condition. Besides, the discrepancy between the experiment and the real situation might influence the perceptions of the recommendation system. To increase credibility by giving the sense of personalized recommendation, we inserted made-up stages that ask for personal information including social media id which is not a common procedure in video recommendation platforms. In addition, we recommended videos to participants with a made-up video list regardless of their preferences which resulted in contrasting outcomes compared to the previous study. To avoid this issue, we could have used a recommendation algorithm that actually recommends personalized videos for each participant. Therefore, researchers should consider a more natural way of making experimental stimuli with ensuring high credibility.

Acknowledgements

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2020-0-01816) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation). This research was also supported by the National Research Foundation of Korea funded by the Korean Government (NRF-2020R1C1C1004324).

References

- Abe, G.; and Richardson, J. 2006. Alarm timing, trust and driver expectation for forward collision warning systems. *Applied ergonomics*, 37(5): 577–586.
- Bagheri, N.; Jamieson, G. A.; et al. 2004. Considering subjective trust and monitoring behavior in assessing automation-induced “complacency.”. *Human performance, situation awareness, and automation: Current research and trends*, 1: 54–59.
- Bedjaoui, M.; Elouali, N.; and Benslimane, S. M. 2018. User time spent between persuasiveness and usability of social networking mobile applications: a case study of Facebook and YouTube. In *Proceedings of the 16th International Conference on Advances in Mobile Computing and Multimedia*, 15–24.
- Bogun, F.; Anh, D.; Kalahasty, G.; Wissner, E.; Serhal, C. B.; Bazzi, R.; Weaver, W. D.; and Schuger, C. 2004. Misdiagnosis of atrial fibrillation and its clinical consequences. *The American journal of medicine*, 117(9): 636–642.
- Cohen, M. S.; Parasuraman, R.; Serfaty, D.; and Andes, R. 1997. Trust in decision aids: A model and a training strategy. Arlington, VA: *Cognitive Technologies, Inc.*
- Dietvorst, B. J.; Simmons, J. P.; and Massey, C. 2015. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1): 114.
- Dzindolet, M. T.; Pierce, L. G.; Beck, H. P.; and Dawe, L. A. 2002. The perceived utility of human and automated aids in a visual detection task. *Human Factors*, 44(1): 79–94.
- Featherman, M. S.; Miyazaki, A. D.; and Sprott, D. E. 2010. Reducing online privacy risk to facilitate e-service adoption: the influence of perceived ease of use and corporate credibility. *Journal of Services Marketing*.
- Fleder, D.; and Hosanagar, K. 2009. Blockbuster culture’s next rise or fall: The effect of recommender systems on sales diversity. *Management Science*, 55(5): 697–712.
- Goddard, K.; Roudsari, A.; and Wyatt, J. C. 2012. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1): 121–127.
- Haim, M.; Graefe, A.; and Brosius, H.-B. 2018. Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital journalism*, 6(3): 330–343.
- Hayes, A. F. 2017. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.
- Higgins, E. T.; Bargh, J. A.; and Lombardi, W. J. 1985. Nature of priming effects on categorization. *Journal of experimental psychology: Learning, Memory, and Cognition*, 11(1): 59.
- Hoff, K. A.; and Bashir, M. 2015. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors*, 57(3): 407–434.
- Hosanagar, K.; Fleder, D.; Lee, D.; and Buja, A. 2014. Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation. *Management Science*, 60(4): 805–823.
- Kim, K. J. 2016. Interacting socially with the Internet of Things (IoT): Effects of source attribution and specialization in human–IoT interaction. *Journal of Computer-Mediated Communication*, 21(6): 420–435.
- Kim, K. J.; Park, E.; and Sundar, S. S. 2013. Caregiving role in human–robot interaction: A study of the mediating effects of perceived benefit and social presence. *Computers in Human Behavior*, 29(4): 1799–1806.
- Koh, Y. J.; and Sundar, S. S. 2010. Heuristic versus systematic processing of specialist versus generalist sources in on-line media. *Human Communication Research*, 36(2): 103–124.
- Layton, C.; Smith, P. J.; and McCoy, C. E. 1994. Design of a cooperative problem-solving system for en-route flight planning: An empirical evaluation. *Human Factors*, 36(1): 94–119.
- Lee, J. D.; and Moray, N. 1994. Trust, self-confidence, and operators’ adaptation to automation. *International journal of human-computer studies*, 40(1): 153–184.
- Mosier, K. L.; Skitka, L. J.; Burdick, M. D.; and Heers, S. T. 1996. Automation bias, accountability, and verification behaviors. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 40, 204–208. SAGE Publications Sage CA: Los Angeles, CA.
- Nie, B.; Zhang, H.; and Liu, Y. 2014. Social interaction based video recommendation: Recommending youtube videos to facebook users. In *Proceedings of 2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 97–102. IEEE.
- Otoni, R.; Cunha, E.; Magno, G.; Bernardina, P.; Meira Jr, W.; and Almeida, V. 2018. Analyzing right-wing youtube channels: Hate, violence and discrimination. In *Proceedings of the 10th ACM Conference on Web Science*, 323–332.
- Parasuraman, R.; and Manzey, D. H. 2010. Complacency and bias in human use of automation: An attentional integration. *Human factors*, 52(3): 381–410.
- Parasuraman, R.; Molloy, R.; and Singh, I. L. 1993. Performance consequences of automation-induced ‘complacency’. *The International Journal of Aviation Psychology*, 3(1): 1–23.
- Pariser, E. 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.
- Singh, I. L.; Molloy, R.; and Parasuraman, R. 1997. Automation-induced monitoring inefficiency: role of display location. *International Journal of Human-Computer Studies*, 46(1): 17–30.

- Skitka, L. J.; Mosier, K. L.; and Burdick, M. 1999. Does automation bias decision-making? *International Journal of Human-Computer Studies*, 51(5): 991–1006.
- Stroud, N. J.; and Lee, J. K. 2013. Perceptions of cable news credibility. *Mass Communication and Society*, 16(1): 67–88.
- Sundar, S. S. 2008. *The MAIN model: A heuristic approach to understanding technology effects on credibility*. MacArthur Foundation Digital Media and Learning Initiative.
- Sundar, S. S. 2020. Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAI). *Journal of Computer-Mediated Communication*, 25(1): 74–88.
- Sundar, S. S.; and Kim, J. 2019. Machine heuristic: When we trust computers more than humans with our personal information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–9.
- Tufekci, Z. 2018. YouTube, the great radicalizer. <https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html>.
- Wickens, C. D.; Clegg, B. A.; Vieane, A. Z.; and Sebok, A. L. 2015. Complacency and automation bias in the use of imperfect automation. *Human factors*, 57(5): 728–739.
- Wolfstadt, J. I.; Gurwitz, J. H.; Field, T. S.; Lee, M.; Kalkar, S.; Wu, W.; and Rochon, P. A. 2008. The effect of computerized physician order entry with clinical decision support on the rates of adverse drug events: a systematic review. *Journal of general internal medicine*, 23(4): 451–458.
- Wölker, A.; and Powell, T. E. 2018. Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. *Journalism*, 1464884918757072.
- Yu, K.; Berkovsky, S.; Conway, D.; Taib, R.; Zhou, J.; and Chen, F. 2018. Do I trust a machine? Differences in user trust based on system performance. In *Human and machine learning*, 245–264. Springer.
- Yuviler-Gavish, N.; and Gopher, D. 2011. Effect of descriptive information and experience on automation reliance. *Human factors*, 53(3): 230–244.