Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Yes-man agent makes user make risky choices

Matsui Tetsuya (tetsuya.matsui@oit.ac.jp)
Osaka Institute of Technology

Short Report

Keywords: human-agent interaction, cognitive bias, decision-making, virtual agent, risky choice

Posted Date: August 16th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3190162/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Yes-man agent makes user make risky choices

Matsui Tetsuya[1*]

[1*]Robotics and Design, Osaka Institute of Technology, Chayamachi, Osaka, 5300013, Japan.

Corresponding author(s). E-mail(s): tetsuya.matsui@oit.ac.jp;

Abstract

In this paper, I focused on subjective probability in decision-making in HAI. Many virtual agents are used to support users' decision-making; however, few prior works have focused on ambiguous situations. I hypothesize that users are more likely to choose ambiguity after interacting with friendly agents than with unfriendly ones. To verify this hypothesis, I conducted an experiment using the Ellsberg problem. In the task, participants chose either a box that definitely contained 30 winning balls or a box that contained an undefined number of winning balls. The independent variable was agents. I used two agents: a friendly yes-man agent and unfriendly agent. As a result, the hypothesis was supported. The participants who interacted with the friendly yes-man agent tended to make risky choices. This demonstrates the peculiarity and risks associated with friendly yes-man agents.

Keywords: human-agent interaction, cognitive bias, decision-making, virtual agent, risky choice

1 Introduction

In the field of human-agent interaction (HAI), many agents have been used to support users' decision-making. Product recommendation virtual agents (PRVAs) are agents that aim to make users buy products [1][2]. Health counseling agents provide health-related advice to users [3][4]. For these purposes, many methods that let users make decisions have been suggested. Braun et al. suggested that an agent should tell lies to support users in the decision-making process [5]. Nudge agents are virtual agents or robots that get users to make a decision or change their behavior on the basis of nudge theorycharters2019empathic[6][7]. Nudge theory suggests that an unaware stimulus

(nudge) can define a person's will and actions; for example, when boarding a train, many people sit in the opposite seat without thinking when the other side of seats is crowded [8][9].

In addition, research on agents that support all aspects of a user's life has been advancing. Virtual coaches are agents that continuously help users lead healthy lives [10][11]. Virtual assistants help users follow their daily time schedule [12][13]. These agents have many opportunities to aid users in decision-making, so long-term human-agent interaction is important. In prior works, long-term interaction agents were designed as "friendly" agents. These agents use self-disclosure to users [14][15], smile or use positive expressions toward users [16][17], and agree with or use adaptive expressions toward users [18][19][20]. However, is there any possibility that these adaptive agents could distort users' decision-making? This is my issue that is verified in this paper. In other words, will yes-man agents have a positive effect on users' decision-making?

2 Aversion to ambiguity

Humans' decision-making is affected by many kinds of biases [21]. Some kinds of biases affect human-human interaction [22][23]. Thus, human-agent interaction also seems to be affected by cognitive biases. However, cognitive biases in human-agent interaction have hardly been researched.

In this paper, I focused on a cognitive bias related to subjective probability. There are many kinds of cognitive biases related to subjective probability that were revealed in prior research. Many people sometimes judge a conditional probability to be higher than a single-event probability (conjunction fallacy)[24]. Many people judge that the probability of an event occurring in subsequent trials will be low if its frequency is high during a given period, even if each trial is independent (gambler's fallacy)[25].

In this paper, I focused on the Ellsberg paradox [26]. This paradox is opposite expected utility theory and regarded as a reflection of people's aversion to ambiguity. In this paper, I introduce a simplified version of the Ellsberg task for one aim: will yes-man agents have a positive effect on users' decision-making?

In the task, there are two blind boxes. There are 90 balls in each box. In box A, there are 30 red balls, 30 blue balls, and 30 yellow balls. In box B, there are 30 red balls and in total 60 blue and yellow balls; the exact numbers of blue and yellow balls are unknown. The question is: Choose one of the boxes and take out one ball without looking inside the box. If you take out a blue ball, you get a prize of 1,000 yen (about 7 dollars). Which box would you choose?

The answer to this question seems to be affected by the person's aversion to ambiguity. For box A, there are definitely 30 blue balls. If you choose box A, the probability that you will get a prize is 30/90. For box B, the number of blue balls is from 0 to 60. If you choose box B, the probability that you will get a prize is from 0/90 to 60/90. Thus, the probability that you will get a prize could be higher than box A, but it could also be lower. The probability that you will get a prize when you choose box B is ambiguous. A stubborn person or person lacking self-confidence will avoid box B.

Incidentally, agents that interact with users in a friendly manner cause users to feel a sense of self-assurance [27][28].

Thus, I hypothesized that:

- H1 A user will more likely choose ambiguity after interacting with a friendly agent than after interacting with an unfriendly agent.
- H2 A user will more likely choose ambiguity after interacting with a friendly agent than without interaction with an agent.

Under these hypothesis, in the above task, a user who interacts with a friendly agent would tend to choose box B. I conducted an experiment to verify this hypothesis.

3 Experiment

In the experiment, I used two types of virtual agents. One agent is named Chiyoko, who always smiles and agrees with participants. This was the friendly agent. The other is named Nichika, who always grimaces and disagrees with participants. This was the unfriendly agent.

In the task, the participants interacted with one of the agents. The experiment design was between participants. The independent value was agent (friendly and unfriendly), and the dependent value was the number of participants who chose ambiguity and the participants' impression of the agent.

In both conditions, the participants interacted with a virtual agent. The agent was our original female model, and the standard was Live2D. She moved and spoke in an experiment movie 1. In both conditions, the agents asked six questions. The first five questions were about small topics, for example, "which do you like, dogs or cats?" In the familiar agent (Chiyoko) condition, the agent always responded by smiling and saying "Excellent! Just like me!" to the participants' answer. In the unfamiliar agent (Nichika) [1]condition, the agent always responded by grimacing and saying "Eh, I can't believe it!" These five questions were to make the agents feel familiar or unfamiliar to the participants.

The sixth (last) question (blind box question) was, "There are two blind boxes. In box A, there are 30 red balls, 30 blue balls, and 30 yellow balls. In box B, there are 30 red balls and in total 60 blue and yellow balls; the exact numbers of blue and yellow balls are unknown. You must choose one of the boxes and take out one ball without looking inside the box. If you take out a blue ball, you get a prize of 1,000 yen (about 7 dollars). Which box would you choose?"

After the interaction, the participants answered questions on their impression of the agents. The questions were as follows.

- Q1: How much did you feel this agent to be trustworthy?
- Q2: How much did you feel this agent to be familiar?
- Q3: How much did you feel this agent to be human-like?
- Q4: How much did you feel this agent to be alive?

[1]Chiyoko and Nichika are the names of characters in the Japanese game "THE IDOLM@STER SHINY COLORS." Chiyoko is a friendly character, and Nichika is rebellious.

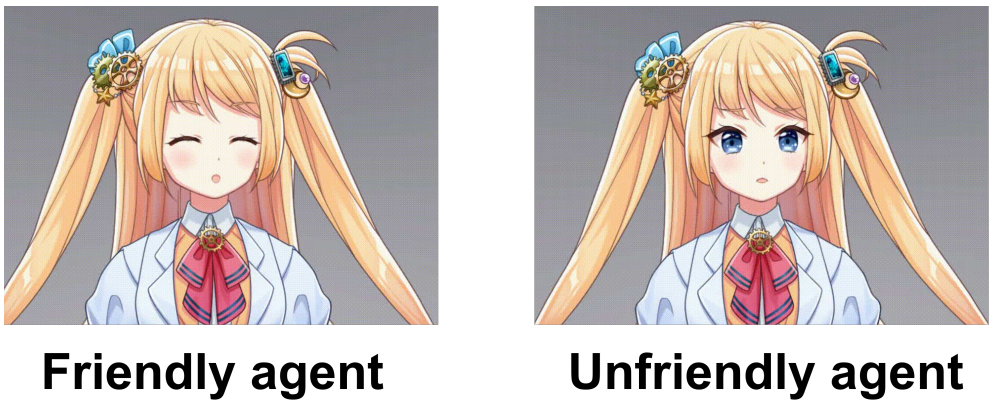


Fig. 1 Agents used in experiment. This figure is an original figure produced by the author(s) for this review article.

The participants answered these questions on a seven-point Likert scale.

I conducted a chi-square test to analyze the answers to the blind box question. The subject of the analysis was the number of participants who chose each answer to the sixth question. Also, I conducted a t-test to analyze the answers to Q1-Q4.

In addition to these two agent conditions, I also set a condition of no interaction with the agent. Agents are not presented with this condition. In this condition, participants answered the same questions as in the familiar agent(chiyoko) and unfamiliar agent(Nichika) conditions. At the end, they answered the blind box question.

In the familiar agent (Chiyoko) condition, there were 161 participants (58 females and 103 males), ranging in age from 20 to 81 years for an average of 45.4 (SD = 11.4). In the unfamiliar agent (Nichika) condition, there were 175 participants (54 females and 121 males), ranging in age from 24 to 64 years for an average of 45.5 (SD = 9.9). This question regarded the level of aversion to ambiguity. In no agent condition, there were 189 participants (54 females, 134 males and 1 other), ranging in age from 22 to 76 years for an average of 47.6 (SD = 10.6).

The experiment was conducted on the web via Yahoo Crowdsourcing[2], and all participants got 50 yen (about 0.35 dollars) as a reward. The credibility of crowd sourcing was proven by Crump et al. [29]. The experiment was conducted with the approval of the ethics committee of the Osaka Institute of Technology.

This experiment was approved by the Ethics Committee of Osaka Institute of Technology(No.2021-25) and I got informed consent via web.

Result

Table 1 shows the number of participants who chose each box in the blind box question. I conducted a chi-square test between each condition.

[2]https://crowdsourcing.yahoo.co.jp/

Table 1 Results for blind box question

agent	number of participants choosing box A	number of participants choosing box B
Chiyoko (familiar)	33	128
Nichika (unfamiliar)	61	114
no agent	58	131

Between familiar agent condition and unfamiliar agent condition, there is a significant difference(χ^2=7.88, p =0.00, ϕ=0.15). As a result of residual analysis, the participants in familiar agent condition were significantly more likely to choose box B than unfamiliar agent condition, and the participants in unfamiliar agent condition were significantly more likely to choose box A than familiar agent condition(residual=2.929).

Between familiar agent condition and no agent condition, there is a significant difference(χ^2=4.18, p =0.04, ϕ=0.11). As a result of residual analysis, the participants in familiar agent condition were significantly more likely to choose box B than no agent condition, and the participants in no agent condition were significantly more likely to choose box A than familiar agent condition(residual=2.166).

Between familiar agent condition and no agent condition, there is a significant difference(χ^2=0.54, p =0.05, ϕ=0.46).

Figure 2 shows the average for Q1-Q4 between familiar agent condition and unfamiliar agent condition. The error bars are the standard deviation, and ∗∗ shows that there was a significant difference (p¡0.01). For Q1, there was a significant difference (p <0.01)(t(175)=-2.5, p =0.00). For Q2, there was significant difference (p <0.01)(t(175)=-6.0, p =0.00). For Q3, there was not a significant difference (t(175)=-1.6, p =0.06). For Q4, there was not a significant difference (t(175)=-1.5, p =0.07).

4 Discussion

Table 1 shows that the difference between the two agents(familiar agent and unfamiliar agent) was significant. In both conditions, box B was chosen more than box A. However, the ratio of participants who chose box B in the friendly agent (Chiyoko) condition was significantly higher than the unfriendly agent (Nichika) condition. This shows the following fact; a friendly agent that always agrees with a user tends to make the user choose ambiguity. A choice of ambiguity may have a higher probability of leading to a prize than box A (when there are 31 or more blue balls in box B); however, it may have a lower probability than box A (when there are 29 or less blue balls in box B). Thus, this is a risky, gambling-like choice. This result supports H 1.

Also, Table 1 shows that the difference between familiar agent condition and no agent condition. The ratio of participants who chose box B in the friendly agent (Chiyoko) condition was significantly higher than the no agent condition. However, there is no significant difference between unfamiliar agent and no agent condition. This result shows that the agent itself did not affect the result. The important thing is

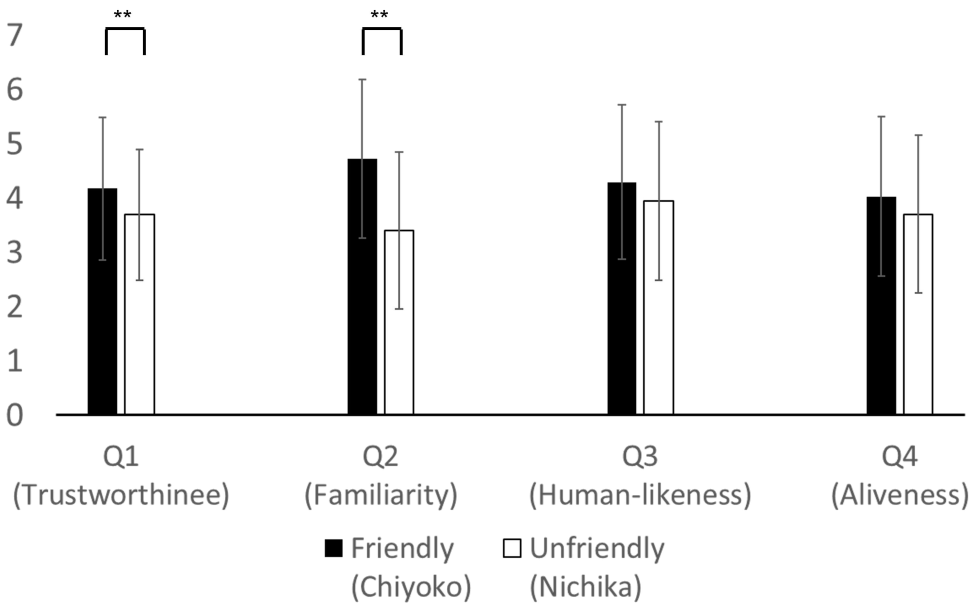


Fig. 2 Average for Q1-Q4 between familiar agent condition and unfamiliar agent condition

whether the agent is friendly or not. With or without an unfamiliar agent, the result is the same. Only the friendly agent affected the participants' decision-making. This result supports H 2.

In short, a friendly yes-man agent can cause a user to make risky and challenging choices. This is ambivalent peculiarity. Such an agent can motivate a user to take on new things; however, they may also cause the user to make a potentially-fatal risky choice. In some contexts, such an agent could be useful, and in other contexts, it may be harmful. In long-term interaction, an agent who always agrees with the user will repeatedly cause the user to make risky choices. This can cause an incident or trouble. Thus, I suggest that such agents must not always agree with the user.

Figure 2 shows that the participants felt the friendly agent to be more trustworthy and familiar than the unfriendly one. This result and Table 1 suggest that the participants felt trustworthiness and familiarity in the agent when they made a risky choice. Otherwise, the agents' human-likeness and aliveness was almost the same between the two agents.

The result of this experiment shows one process by which a users' subjective probability is affected by an agent. A friendly agent increases a user's self-assurance [27][28], and a high self-assurance causes a user to make risky, gambling-like choices. A user who interacts with a friendly yes-man agent tends to interpret things in a positive light. This shows the risky aspect of friendly yes-man agents.

5 Conclusion

In this paper, I focused on subjective probability in decision-making in HAI. I hypothesized that a user will more likely choose ambiguity after interacting with a friendly agent than after interacting with an unfriendly one. To verify this hypothesis, I conducted an experiment using the Ellsberg problem. I used two agents: a friendly yes-man agent and an unfriendly agent. As a result, the hypothesis was supported. The participants who interacted with the friendly yes-man agent tended to make risky choices. This demonstrates the peculiarity and risks associated with friendly yes-man agents.

This result suggests a problem that occurs when a user interacts with agents over a long period of time and suggests a new aspect in designing virtual agents.

6 Data availability statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

[1] Qiu, L., Benbasat, I.: Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. Journal of management information systems **25**(4), 145–182 (2009)

[2] Matsui, T., Yamada, S.: Designing trustworthy product recommendation virtual agents operating positive emotion and having copious amount of knowledge. Frontiers in psychology **10**, 675 (2019)

[3] Bickmore, T.W., Pfeifer, L.M., Paasche-Orlow, M.K.: Health document explanation by virtual agents. In: International Workshop on Intelligent Virtual Agents, pp. 183–196 (2007)

[4] Olafsson, S., O'Leary, T.K., Bickmore, T.W.: Motivating health behavior change with humorous virtual agents. In: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, pp. 1–8 (2020)

[5] Braun, D., Bhat, M., Biesdorf, A., Matthes, F.: Would you lie to me bot? supporting decision-making processes with deceiving virtual agents. Procedia Computer Science **177**, 587–592 (2020)

[6] Hang, C., Ono, T., Yamada, S.: Designing nudge agents that promote human altruism. In: International Conference on Social Robotics, pp. 375–385 (2021)

[7] Stirapongsasuti, S., Thonglek, K., Misaki, S., Nakamura, Y., Yasumoto, K.: Insha: Intelligent nudging system for hand hygiene awareness. In: Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents, pp. 183–190 (2021)

[8] Abdukadirov, S.: Nudge Theory in Action: Behavioral Design in Policy and Markets, (2016)

[9] Arno, A., Thomas, S.: The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis. BMC public health **16**(1), 1–11 (2016)

[10] Akker, H., Klaassen, R., Nijholt, A.: Virtual Coaches for Healthy Lifestyle. (2016)

[11] Kulyk, O., Akker, R., Klaassen, R., Gemert-Pijnen, L.: Personalized virtual coaching for lifestyle support: principles for design and evaluation. International journal on advances in life sciences **6**(3-4), 300–309 (2014)

[12] Tulshan, A.S., Dhage, S.N.: Survey on virtual assistant: Google assistant, siri, cortana, alexa. In: International Symposium on Signal Processing and Intelligent Recognition Systems, pp. 190–201 (2018)

[13] Gupta, K., Hajika, R., Pai, Y.S., Duenser, A., Lochner, M., Billinghurst, M.: Measuring human trust in a virtual assistant using physiological sensing in virtual reality. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 756–765 (2020)

[14] Kang, S.-H., Gratch, J., Sidner, C., Artstein, R., Huang, L., Morency, L.-P.: Towards building a virtual counselor: modeling nonverbal behavior during intimate self-disclosure. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1, pp. 63–70 (2012)

[15] Burger, F., Broekens, J., Neerincx, M.A.: Fostering relatedness between children and virtual agents through reciprocal self-disclosure. In: Benelux Conference on Artificial Intelligence, pp. 137–154 (2016)

[16] Numata, T., Sato, H., Asa, Y., Koike, T., Miyata, K., Nakagawa, E., Sumiya, M., Sadato, N.: Achieving affective human–virtual agent communication by enabling virtual agents to imitate positive expressions. Scientific reports **10**(1), 1–11 (2020)

[17] Ochs, M., Niewiadomski, R., Pelachaud, C.: How a virtual agent should smile? In: International Conference on Intelligent Virtual Agents, pp. 427–440 (2010)

[18] Zhang, H., Fricker, D., Smith, T.G., Yu, C.: Real-time adaptive behaviors in multimodal human-avatar interactions. In: International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction, pp. 1–8 (2010)

[19] Niewiadomski, R., Hofmann, J., Urbain, J., Platt, T., Wagner, J., Bilal, P., Ito, T., Jonker, C., Gini, M., Shehory, O.: Laugh-aware virtual agent and its impact on user amusement (2013)

[20] Ben Youssef, A., Chollet, M., Jones, H., Sabouret, N., Pelachaud, C., Ochs, M.: Towards a socially adaptive virtual agent. In: International Conference on Intelligent Virtual Agents, pp. 3–16 (2015)

[21] Blanco, F.: Cognitive bias. Encyclopedia of animal cognition and behavior **1**(6) (2017)

[22] Lean Keng, S., AlQudah, H.N.I.: Assessment of cognitive bias in decision-making and leadership styles among critical care nurses: a mixed methods study. Journal of advanced nursing **73**(2), 465–481 (2017)

[23] Haselton, M., Galperin, A.: Error management and the evolution of cognitive bias. Soc. Think. Interpers. Behav. 45 **63** (2012)

[24] Tentori, K., Bonini, N., Osherson, D.: The conjunction fallacy: A misunderstanding about conjunction? Cognitive Science **28**(3), 467–477 (2004)

[25] Barron, G., Leider, S.: The role of experience in the gambler's fallacy. Journal of Behavioral Decision Making **23**(1), 117–129 (2010)

[26] Segal, U.: The ellsberg paradox and risk aversion: An anticipated utility approach. International Economic Review, 175–202 (1987)

[27] Liew, T., Azan Mat Zin, N., Sahari, N., Tan, S.-M.: The effects of a pedagogical agent's smiling expression on the learner's emotions and motivation in a virtual learning environment. International Review of Research in Open and Distributed Learning: IRRODL **17**(5), 248–266 (2016)

[28] Bee, N., André, E., Tober, S.: Breaking the ice in human-agent communication: Eye-gaze based initiation of contact with an embodied conversational agent. In: International Workshop on Intelligent Virtual Agents, pp. 229–242 (2009)

[29] Crump, M.J., McDonnell, J.V., Gureckis, T.M.: Evaluating amazon's mechanical turk as a tool for experimental behavioral research. PloS one **8**(3), 57410 (2013)