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.

1+1/ = 2: Appearance and utterance to determine the virtual agent's intellectual impression

Tetsuya Matsui

matsui.tetsuya@kagawa-u.ac.jp

Kagawa University

Kazuto Sasai

Ibaraki University

Iori Tani

Kobe University

Research Article

Keywords: Virtual agents, anthropomorphic agents, intelligence perceived, trustworthy agents, human-agent interaction

Posted Date: October 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10799307/v1

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

Additional Declarations: No competing interests reported.

Abstract

Virtual agents are beginning to find practical applications in many aspects of society. One way to increase user trust in virtual agents is to make them perceived as more intelligent. The appearance and speech of virtual agents that users perceive as intelligent have each been studied alone. However, whether these can be combined to create agents that are perceived as more intelligent has not yet been studied. In this paper, a two-factor, two-level experiment was conducted to examine the appearance and speech of virtual agents at two levels, “intelligent” and “unintelligent.” The results showed that “agents who look intelligent but speak in an unintelligent manner” and “agents who look unintelligent but speak intelligently” were rated as more intelligent and trusted by users than “agents who look intelligent and speak intelligently.”

1 Introduction

Anthropomorphic virtual agents do not have physical bodies but look like humans, animals, robots, and other similar beings. They are used in various parts of society: product recommendation (Qiu and Benbasat 2009; Ben Saad and Choura 2023), education(Martha and Santoso 2019; Grivokostopoulou et al. 2020) and counseling (Olafsson et al. 2020). HAI(human-agent interaction) is a research field that focuses on the interaction design between a user(human) and an artificial agent(virtual agents, robots, and AIs). HAI studies have shown that it is important for agents to be trusted by their users. According to research on virtual agent trustworthiness, one way to increase user perceptions of virtual agent trustworthiness is to make virtual agents perceived as intelligent. As previous works suggested, robots’ performance and intelligence are significant factors in their perceived trustworthiness (Hancock et al. 2011; Billings et al. 2012). The same is applies for virtual agents, where users trust intelligent agents more (Glikson and Woolley 2020; Matsui and Yamada 2019; Tan and Liew 2020).

The user judges whether a robot or agent is intelligent based on utterance and appearance. Users judge virtual agents that speak with specialized knowledge to be intelligent (Matsui and Yamada 2019). On the other hand, users also judge intelligence from appearance (Sun and Botev, 2021). Chita et al. showed that robots’ gender affected the emotional intelligence perceived by users (Sun and Botev 2021). Walters et al. showed that a robot with a head was perceived as more intelligent than one without a head (Walters et al. 2009). These studies suggest that users judge agents’ intelligence from appearance.

The appearance of a virtual agent is related to factors other than intelligence as perceived by the user. Komatsu et al. showed that users could infer some degree of personality from the design of the virtual agent’s eyes (Komatsu et al. 2016). In humanrobot interaction, it has also been reported that a robot’s appearance, especially its facial design, influences the user’s evaluation of it (Haring et al. 2016; Walters et al. 2009; Crowell et al. 2019).

In summary, users judge an agent’s intelligence based on “the agent’s appearance” and “the agent’s utterance.” It may be valid to conclude that combining “intelligent-looking appearance” and “intelligent utterance” results in the most intelligent perceived agent. However, it is not yet fully understood how

virtual agents' appearance and utterance relate to their performance and personality evaluation. We do not know whether combining "intelligent" elements in the abovementioned ways can improve agents' performance.

Matsui and Gunji showed that the combination of agent elements in user cognition is not arbitrary but has certain constraints (Matsui and Gunji 2021). Matsui also showed that whether a user perceives a female agent as sexual or not is not determined solely by the agent's appearance (Matsui 2022) but also depends on the context in which the agent is used, not just her clothing. These studies indicate that the agent's impression comprises complex processes such as context and interactions between elements, and combining "intelligent appearance" and "intelligent speech" may not work as expected. Therefore, this study explores how these factors can be combined to design agents perceived as more intelligent and trustworthy.

Behavioral economics has shown that human decisions are not entirely rational. The Linda problem (Morier and Borgida 1984) and the guppy effect (Smith and Osherson 1984) are prime examples. The Linda problem is related to a cognitive paradox called the "conjunction fallacy." If considered completely mathematically, the probability that a person satisfies conditions A and B simultaneously is clearly lower than the probability that he or she satisfies condition A alone. Nevertheless, depending on the way of the question, the experiment participants may conclude that the probability of satisfying A and B is higher than the probability of satisfying A alone. The guppy effect is another issue related to a similar cognitive paradox. When asked to name a typical fish or pet, most people do not mention guppies. Nevertheless, many people name guppies when asked to name a typical pet fish, and the probability of naming guppies in this condition is much larger than the joint probability of naming guppies in individual questions. Quantum cogntion studies has used quantum logic to model seemingly irrational decision-making, such as above (Ludwin-Peery et al. 2020; Aerts et al. 2012; Ozawa and Khrennikov 2019).

Conventional HAI studies have assumed that we can freely combine the agent elements concerning intelligence (Matsui and Yamada 2019; Sun and Botev 2021; Chita-Tegmark et al. 2019; Walters et al. 2009). This assumption implies that the overall impression of the virtual agent is determined by simply adding or subtracting the component impressions. For example, this addition/subtraction assumes that we can design a more intelligent perceived agent by combining "appearance" and "utterance" that the user perceives. The connection between "looking intelligent" and "speaking intelligently" is usually perceived as self-consistent and self-evident. Adaptive gap theory also assumes that the user expects artificial agents (e.g., robots) with an intelligent appearance to speak intelligently (Komatsu and Yamada 2011). However, it is unclear whether a linear mathematical equation can model virtual agents' impressions since they are determined by a complex process that depends on the interaction of multiple factors and the context in which they are used.

Therefore, we experimented to determine how the combination of an agent's appearance and speech style affects the agent's impression. Specifically, we prepared two levels, intelligent or unintelligent in

two factors, appearance and utterance, and examined the effects of each combination on the user's evaluation of the agent. We tested the following hypothesis by experiment;

H: An agent that combines "intelligent appearance and intelligent utterance" perceived as more intelligent than others.

Experimental results rejected the hypothesis. We revealed that users perceive agents who "look intelligent but speak in an unintelligent manner" and "look unintelligent but speak in an intelligent manner" as more intelligent.

To explain this confusing result, we focused on the recently proposed natural-born intelligence theory, which can comprehensively explain seemingly irrational human decisions that are explained by quantum cognition theory (Gunji and Nakamura 2022). This theory also supports the edge of chaos argument (Gunji and Uragami 2021) and the connection between the free energy principle and quantum cognition (Gunji et al. 2022). This theory allows us to interpret our experimental results successfully.

In the next section, we describe our experiments to test the hypothesis. First, we describe our preliminary experiments to define "intelligent appearance" and "unintelligent appearance." Next, the main two-factor, two-level experiment, in which appearance and utterance were used as factors and a level of "intelligent" or "unintelligent" was assigned to each, is described. We conducted all experiments on the web using a crowdsourcing service and analyzed the results using ANOVA. In the discussion section, we conceptually introduce the natural-born intelligence theory in the context of our experiment and show that this theory can explain the experimental results well. Moreover, we will discuss prospects for future research.

2 Materials & Methods

The experiments proceeded as follows. First, we conducted preliminary experiments to define "intelligent appearance" and "unintelligent appearance." We constructed five female virtual agents. The only difference in appearance between these agents is their hairstyle; otherwise, all components are the same. The preliminary experiments were pairwise comparisons for all combinations using these five agents. Experimental participants were shown two agents and asked, "Which agent do you think is more intelligent?" The preliminary experiments allowed us to determine the "intelligent appearance" and the "unintelligent appearance." The main experiments were a two-factor, two-level experiment combining appearance and speech text. The "intelligent text" and "unintelligent text" were taken from prior studies. In the experiments, virtual agents made travel recommendations in the movie. Participants answered questions about the agent's intelligence and trustworthiness perceived after watching the movie. All experiments were conducted with the approval of the Osaka Institute of Technology ethics committee.

2.1 Preliminary experiment: agents' appearance

We used “agents’ appearance” as the first factor in the main experiments. So, we had to define the agents’ “intelligent appearance” and “unintelligent appearance” through preliminary experiments. We constructed five different-looking agents using VRoid Studio, released by Pixiv Inc. . All agents differed only in hairstyles; all other elements were identical.

In the preliminary experiment, Participants were shown pictures of a pair of the five agents and asked which agent looked more intelligent. The participants were required to make pairwise comparisons for all combinations. After evaluating all combinations, participants responded with their gender and age to complete the experiment. In the preliminary experiment, there were 86 participants (18 females and 68 males), ranging in age from 22 to 70 years, for an average of 44.9 (SD = 10.3). The preliminary experiment was conducted on the web via Yahoo Crowdsourcing , and all participants got 20 yen (about 0.14 dollars) as a reward. Many studies have validated the credibility of crowdsourcing (Crump et al. 2013; Giudice 2010; Huang and Sundar 2022).

We analyzed our results using Thurston’s pairwise comparison method. Figure 1 shows the result. As a result, Agent A (short hair) was rated as the most intelligent looking, while Agent B (pigtails) was rated as the least intelligent looking. Therefore, we defined Agent A as “an agent with an intelligent appearance” and Agent B as “an agent with a unintelligent appearance.” We used these two agents in the main experiment.

2.2 Speech text

The second factor is “speech text.” We used two types of texts: Intelligent and unintelligent style. We referred to these texts from previous studies, Matsui and Yamada(2019) (Matsui and Yamada 2019). In this previous study, the authors prepared intelligentstyle texts containing expert knowledge and unintelligent-style texts that did not contain them and experimented with their implementation in virtual agents. We used the intelligent and unintelligent styled texts used in this previous study.

2.3 Evaluation scales

In the main experiment, we needed to measure the agent’s intelligence and trustworthiness perceived by participants to test our hypothesis. We adopted the scale used in previous studies to evaluate intelligence (Matsui and Yamada 2019). There are several other measure scales of user-perceived intelligence for robots and virtual agents in prior research(Bartneck et al. 2009; Chollet 2019). This scale has been validated in previous studies and is advantageous because it consists of a few questionnaires, reducing the participants’ burden.

To measure trustworthiness, we prepared the following three questions;

Q1 How credible did you find this agent?

Q2 How persuasive did you find this agent?

Q3 To what extent did you want to hear what this agent had to say?

Participants answered each question on a 7-point Likert scale ranging from 1 (not at all) to 7 (very much). The total was calculated and used as the ”trustworthiness” scale.

2.4 Main experiment

The main experiment was conducted on a two-factor, two-level. Two factors are the virtual agent’s appearance and the utterance’s text. The levels are “intelligent” and “unintelligent” for both factors. These are as defined in the previous subsections. In the main experiment, a virtual agent recommended a trip to Gifu Castle, one of Japan’s tourist attractions in the video. This task design is the same as a previous study’s (Matsui and Yamada 2019). After watching the video, participants answered the questionnaire about perceived agents’ intelligence and trustworthiness, as mentioned in the previous section. Finally, participants completed the experiment by answering their gender and age. For the results, we performed a two-factor analysis of variance.

The main experiment was conducted on the web via Yahoo Crowdsourcing, just as in the preliminary experiment, and all participants got 50 yen (about 0.35 dollars) as a reward. Table 1 shows the conditions in the main experiment. For condition 1, there were 63 participants (21 females, 41 males, and one other), ranging in age from 21 to 50 years for an average of 39.6 (SD = 8.0). For condition 2, there were 57 participants (21 females, 36 males, and one other), ranging in age from 20 to 50 years, for an average of 41.4 (SD = 6.7). For condition 3, there were 53 participants (19 females and 34 males), ranging in age from 23 to 50 years, for an average of 42.3 (SD = 6.7). For condition 4, there were 54 participants (14 females, 39 males, and one other), ranging in age from 21 to 50 years, for an average of 41.2 (SD = 6.6). Figure 2 is the snapshot of the movie used in the main experiment.

3 Results

Table 2 shows the averages and SDs of scores for each question. We performed a two-factor analysis of variance on the results of intelligence and trustworthiness, respectively. The top of Table 3 shows the results of the three-way ANOVA for intelligence. A statistically significant interaction exists between appearance × text ($p < 0.05$). The bottom of Table 3 shows the results of a simple interaction test between appearance × text. Significant differences existed for the simple main effect in appearance when the text was unintelligent ($p < 0.01$). Figure 3 shows this interaction. When text is unintelligent, the mean score of intelligence is higher when appearance is at an intelligent level than when appearance is unintelligent.

Table 1
Conditions in experiment

Condition	appearance	text
Condition 1	intelligent	intelligent
Condition 2	intelligent	unintelligent
Condition 3	unintelligent	intelligent
Condition 4	unintelligent	unintelligent

The top of Table 4 shows the results of the three-way ANOVA for trustworthiness. There is a statistically significant main effect of appearance($p < 0.05$) and interactions between appearance × text ($p < 0.05$). The bottom of Table 4 shows the results of a simple interaction test for the interactions between appearance × text. Significant differences existed for the simple main effect in appearance when the text was unintelligent ($p < 0.01$). Figure 4 shows this interaction. When text is at an unintelligent level, the mean score of trustworthiness is higher when appearance is at an intelligent level than when appearance is at an unintelligent level.

Table 2
The averages and SDs of scores for each scale.

Condition	intelligence	trustworthiness
Condition 1	21.3(5.37)	12.5(3.5)
Condition 2	23.1(4.35)	13.8(2.92)
Condition 3	22.2(5.29)	12.9(2.96)
Condition 4	20.2(5.27)	11.8(3.67)

We tested whether the participants' gender played a role in the results of this experiment. We conducted female vs. male t-tests for gender differences in " trustworthiness" and" intelligence" scores across all conditions. The results were $p > 0.1$ for all tests and concluded that there was no significant effect of participant gender on this experiment. Table 5 shows these results.

4 Discussion

Table 3 and Fig. 3 show that there is a significant interaction of appearance × text. When the mean values were compared, the [intelligent appearance, unintelligent text] and [non-intelligent appearance, intelligent text] conditions had higher values for intelligence perceived than the [intelligent appearance, intelligent text] condition. This result shows that it is not the combination of elements that perceive intelligence but rather the inclusion of unintelligent elements that creates an agent with an overall impression of intelligence. It contradicts the traditional assumption that a simple summation of elements' impressions can express agent impressions.

Table 3
The results of ANOVA for scores of intelligence. There is a significant main effect of interaction.

source	**F**	**p**
appearance	3.45	0.06
text	0.02	0.90
interaction ∗∗	6.78	0.01
sub-effect test (simple main effect) effect	F	p
appearance (intelligent text)	0.28	0.60
appearance (unintelligent text)	9.98	0.00
∗∗ text (intelligent appearance)	3.09	0.08
text (unintelligent appearance)	3.72	0.06

Table 4
The results of ANOVA for the scale of trustworthiness. There are significant main effects of appearance and interaction.

source	**F**	**p**	
appearance	4.61	0.03	∗
text	0.11	0.74	
interaction	5.38	0.02	∗∗
sub-effect test (simple main effect) effect	F	p	
appearance (intelligent text)	0.02	0.90	
appearance (unintelligent text)	9.98	0.00	∗∗
text (intelligent appearance)	3.53	0.06	
text (unintelligent appearance)	1.97	0.16	

Table 4 and Fig. 4 show that agents' perceived trustworthiness has a similar result. This result is consistent with previous research that agent trustworthiness is influenced by the agent's perceived intelligence (Glikson and Woolley 2020; Matsui and Yamada 2019; Tan and Liew 2020).

As a result, the [intelligent appearance, unintelligent text] and [non-intelligent appearance, intelligent text] agents had higher values for trustworthiness perceived than the [intelligent appearance, intelligent text]

agents. The results show that combining intelligent elements does not create an agent perceived as intelligent, and including non-intellectual elements in appearance or speech is more effective. These results force a rethinking of traditional agent design methods in HAI.

The traditional HAI framework could not explain this result. However, this result is consistent with findings in behavioral economics that humans do not always make rational decisions (Camerer and Loewenstein 2004; Aerts et al. 2012). To explain this irrational result, we refer to the recently proposed natural-born intelligence theory (Gunji and Nakamura 2022).

We describe the natural-born intelligence theory in the context of our experiment. In this theory, we first consider a positive antinomy of the form “A and not A.” It corresponds to the agent with conflict in the intelligence exhibited by its appearance and speech in our experiment. The change of context brings “outside” of the existing context by changing the positive antinomy to the negative antinomy of the form “neither A nor not A.” What this “outside” is cannot be predicted, but the intelligence that accepts it is called natural-born intelligence.

In our experiment, the interaction between the participants and the agent with conflict changed the positive antinomy structure, and the whole interaction was considered to constitute natural-born intelligence. In other words, the conflict between the impression of intelligence given by the agent’s appearance and utterance would evoke an impression of the agent’s personality, familiarity, or latent intellectual impression beyond that given by the individual elements. The fact that the agents used in the experiment had a human-like appearance also helped the participants recall the agents’ personalities, which were not perceived solely from their appearance or utterances.

This may be partly due to the fact that the videos used in the experiment were in the commercial context of travel recommendations. Commercial recommendations are generally provided in an informative, i.e., intelligent manner. An agent with the conflict between appearance and speech regarding intelligence would have altered this usual commercial context and evoked a sense of trustworthiness and familiarity. However, we should note that the same results were not obtained for agents with unintelligent appearances and utterances; they were rated lower for intelligence and trustworthiness. This suggests that the manifestation of contradictions in the agents themselves is necessary for contextual change.

Figure 5 (A) illustrates the concept of natural-born intelligence in this study. The upper left and lower right circles represent the opposing concepts that constitute antinomy. The context of interaction (bottom right circle) acts on these, and the antinomy structure and context itself are altered, resulting in the emergence of an “outside” of the existing context. In the case of this experiment, the “outside” was the agent’s personality and sense of familiarity. On the other hand, Fig. 5 (B) shows the case where no inconsistency exists in the agent. No antinomy is formed in this case, and the interaction context is unchanged. Therefore, the “outside” of the context does not appear, and the participants have evaluated the agent only based on what was presented.

Table 5 Result of female vs. male t-tests for the scale of trustworthiness and intelligence. There are no significant differences.

condition	scalet	p
condition 1	**trustworthiness0.81**	**0.21**
condition 1	intelligence0.06	0.47
condition 2	trustworthiness0.88	0.19
condition 2	intelligence0.60	0.28
condition 3	trustworthiness0.54	0.30
condition 3	intelligence1.04	0.15
condition 4	trustworthiness-0.31	0.38
condition 4	intelligence-0.75	0.23

Finally, we discuss the limitations of this study. This study focuses only on “appearance” and “utterance” among the many factors of virtual agents. It is unclear whether similar results can be obtained when combining other factors. In addition, we used only female virtual agents looking like humans in this study. Suppose the agent is male or has a different appearance than a human, such as a robot or animal. What emerges as “outside” of the context may be different than for a female agent, even if the same conflict exists between their appearance and utterance. Furthermore, we need to consider the case where the agent is used in a context other than commercial recommendations.

5 Conclusion

In this study, we tested by experiment whether combining elements with specific characteristics in the design of a virtual agent enhances those characteristics. To test this hypothesis, we focused on “whether or not the agent looks intelligent” and “whether or not the utterance is intelligent.” In a preliminary experiment, we identified appearances that were perceived as intelligent and unintelligent. For intelligent and unintelligent utterances, we used sentences from previous studies and combined them in a two-factor, two-level experiment. The results showed that the agents appeared more intelligent and trusted the combination of “intelligent appearance” and “unintelligent utterance” or “unintelligent appearance” and “intelligent utterance” than the combination of “intelligent appearance” and “intelligent utterance.” This result cannot be explained in the conventional HAI framework but can be explained using naturalborn intelligence theory, which consistently explains human irrational decision-making. These results force a reconsideration of the traditional virtual agent design method.

Declarations

Consent to Participate

Informed consent was obtained electronically from all participants.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Human Ethics and Consent to Participate declarations

The study was conducted with the approval of the Research Ethics Committee of Osaka Institute of Technology(2023-38). The study was conducted in accordance with the Declaration of Helsinki.

Consent to Publish

Not applicable

Funding

No funding

Author Contribution

T.M., K.S., and I.T. jointly designed the experimental methodology. T.M. took the primary role in conducting the experiments. T.M. and I.T. performed the data analysis. All authors contributed to the interpretation and discussion of the results. T.M. wrote the initial draft of the manuscript, and all authors reviewed and approved the final manuscript.

Data availability

The data that support the findings of this study are available from the authors upon reasonable request

References

1. Aerts D, Broekaert J, Gabora L et al. (2012) The guppy effect as interference. In: Quantum Interaction: 6th International Symposium, QI 2012, Paris, France, June 27–29, 2012, Revised Selected Papers 6, pp 36–47.

2. Bartneck C, Kuli´c D, Croft E, et al. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int J Social Robot. 2009;1:71–81.
3. Ben Saad S, Choura F. Towards better interaction between salespeople and consumers: The role of virtual recommendation agent. Eur J Mark. 2023;57(3):858–903. https://doi.org/10.1108/EJM-11-2021-0892.
4. Billings DR, Schaefer KE, Chen JY et al. (2012) Human-robot interaction: developing trust in robots. In: Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, pp 109–110.
5. Camerer CF, Loewenstein G. Behavioral economics: Past, present, future. Adv Behav Econ. 2004;1:3–51.
6. Chita-Tegmark M, Lohani M, Scheutz M. (2019) Gender effects in perceptions of robots and humans with varying emotional intelligence. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp 230–238.
7. Chollet F. (2019) On the measure of intelligence. arXiv preprint arXiv:191101547.
8. Crowell CR, Deska JC, Villano M, et al. Anthropomorphism of robots: study of appearance and agency. JMIR Hum factors. 2019;6(2):e12629.
9. Crump MJ, McDonnell JV, Gureckis TM. Evaluating amazon's mechanical turk as a tool for experimental behavioral research. PLoS ONE. 2013;8(3):e57410.
10. Giudice KD. (2010) Crowdsourcing credibility: The impact of audience feedback on web page credibility. Proceedings of the American Society for Information Science and Technology 47(1):1–9.
11. Glikson E, Woolley AW. Human trust in artificial intelligence: Review of empirical research. Acad Manag Ann. 2020;14(2):627–60.
12. Grivokostopoulou F, Kovas K, Perikos I. The effectiveness of embodied pedagogical agents and their impact on students learning in virtual worlds. Appl Sci. 2020;10(5):1739.
13. Gunji YP, Nakamura K. Psychological origin of quantum logic: An orthomodular lattice derived from natural-born intelligence without hilbert space. BioSystems. 2022;215:104649.
14. Gunji YP, Uragami D. Computational power of asynchronously tuned automata enhancing the unfolded edge of chaos. Entropy. 2021;23(11):1376.
15. Gunji YP, Shinohara S, Basios V. Connecting the free energy principle with quantum cognition. Front Neurorobotics. 2022;16:910161.
16. Hancock PA, Billings DR, Schaefer KE, et al. A meta-analysis of factors affecting trust in human-robot interaction. Hum Factors. 2011;53(5):517–27.
17. Haring KS, Silvera-Tawil D, Watanabe K et al. (2016) The influence of robot appearance and interactive ability in hri: a cross-cultural study. In: Social Robotics: 8th International Conference, ICSR 2016, Kansas City, MO, USA, November 1–3, 2016 Proceedings 8, Springer, pp 392–401.
18. Huang Y, Sundar SS. Do we trust the crowd? effects of crowdsourcing on perceived credibility of online health information. Health Commun. 2022;37(1):93–102.

19. Komatsu T, Yamada S. Adaptation gap hypothesis: How differences between users' expected and perceived agent functions affect their subjective impression. J Systemics Cybernetics Inf. 2011;9(1):67–74.
20. Komatsu T, Kuramoro I, Sawai D. (2016) Can different eye designs for anthropomorphic manga characters inform users of different functions of anthropomorphized systems? In: Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology, pp 1–6.
21. Ludwin-Peery E, Bramley NR, Davis E, et al. Broken physics: A conjunctionfallacy effect in intuitive physical reasoning. Psychol Sci. 2020;31(12):1602–11.
22. Martha ASD, Santoso HB. The design and impact of the pedagogical agent: A systematic literature review. J Educators Online. 2019;16(1):n1.
23. Matsui T. (2022) Moe-phobia: Effect of users' gender on perceived sexuality and likability toward manga-like virtual agents. Front Psychol 13.
24. Matsui T, Gunji YP. Experimental disproof of a manga character construction model. Symmetry. 2021;13(5):838.
25. Matsui T, Yamada S. Designing trustworthy product recommendation virtual agents operating positive emotion and having copious amount of knowledge. Front Psychol. 2019;10:675.
26. Morier DM, Borgida E. The conjunction fallacy: A task specific phenomenon? Pers Soc Psychol Bull. 1984;10(2):243–52.
27. Olafsson S, Wallace BC, Bickmore TW. (2020) Towards a computational framework for automating substance use counseling with virtual agents. In: AAMAS, pp 966–974.
28. Ozawa M, Khrennikov A. Application of theory of quantum instruments to psychology: Combination of question order effect with response replicability effect. Entropy. 2019;22(1):37.
29. Qiu L, Benbasat I. Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. J Manage Inform Syst. 2009;25(4):145–82.
30. Smith EE, Osherson DN. Conceptual combination with prototype concepts. Cogn Sci. 1984;8(4):337–61.
31. Sun N, Botev J. Intelligent autonomous agents and trust in virtual reality. Computers Hum Behav Rep. 2021;4:100146.
32. Tan SM, Liew TW. Designing embodied virtual agents as product specialists in a multi-product category e-commerce: The roles of source credibility and social presence. Int J Human–Computer Interact. 2020;36(12):1136–49.
33. Walters ML, Koay KL, Syrdal DS, et al. Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials. Procs of New; 2009. Frontiers. in Human-Robot Interaction.

Footnotes

1. https://vroid.com/studio
2. https://crowdsourcing.yahoo.co.jp/

Figures

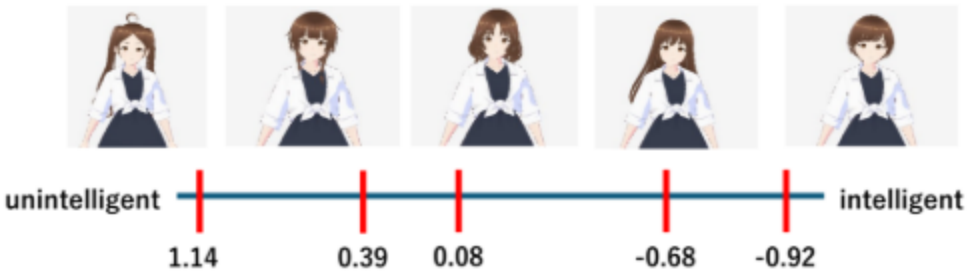

Figure 1

In a preliminary experiment, five agents were used, and a value indicating intelligence for appearance was obtained through a pairwise comparison. The smaller this value, the more intelligent the subject rated the agent. These agents had the same appearance except for their hairstyles.

Figure 2

Snapshot of the video used in the main experiment. This agent is Agent A (intelligent looking appearance).

Figure 3

Averages of intelligence. Error bar means SDs. There is an interaction.

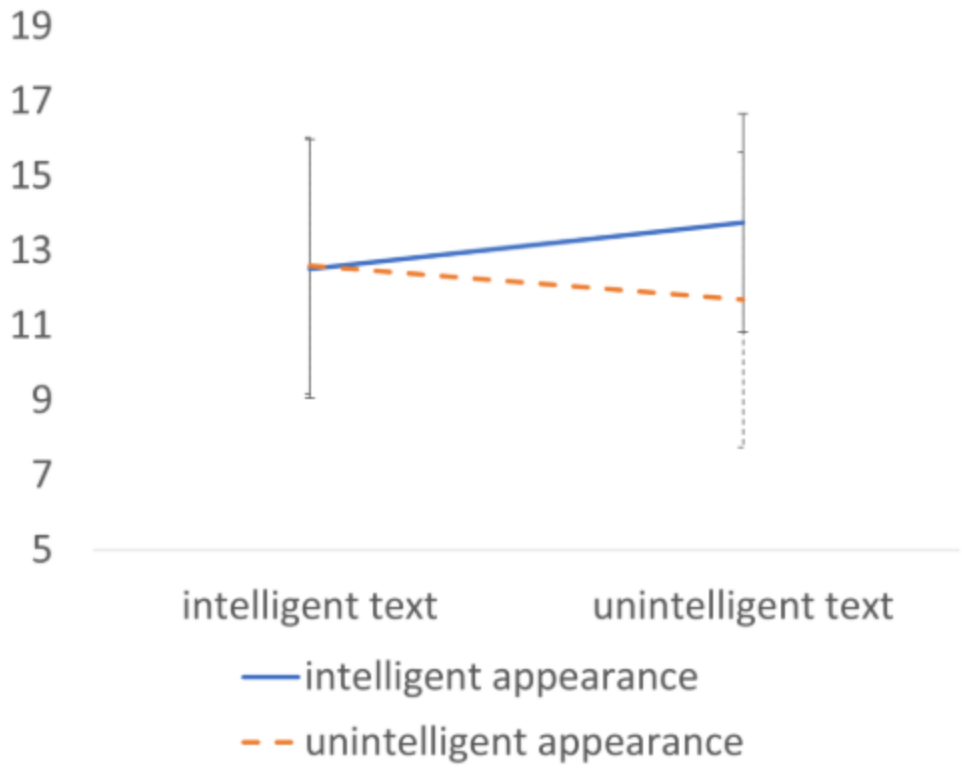

Figure 4

Averages of trustworthiness. Error bar means SDs. There is an interaction.

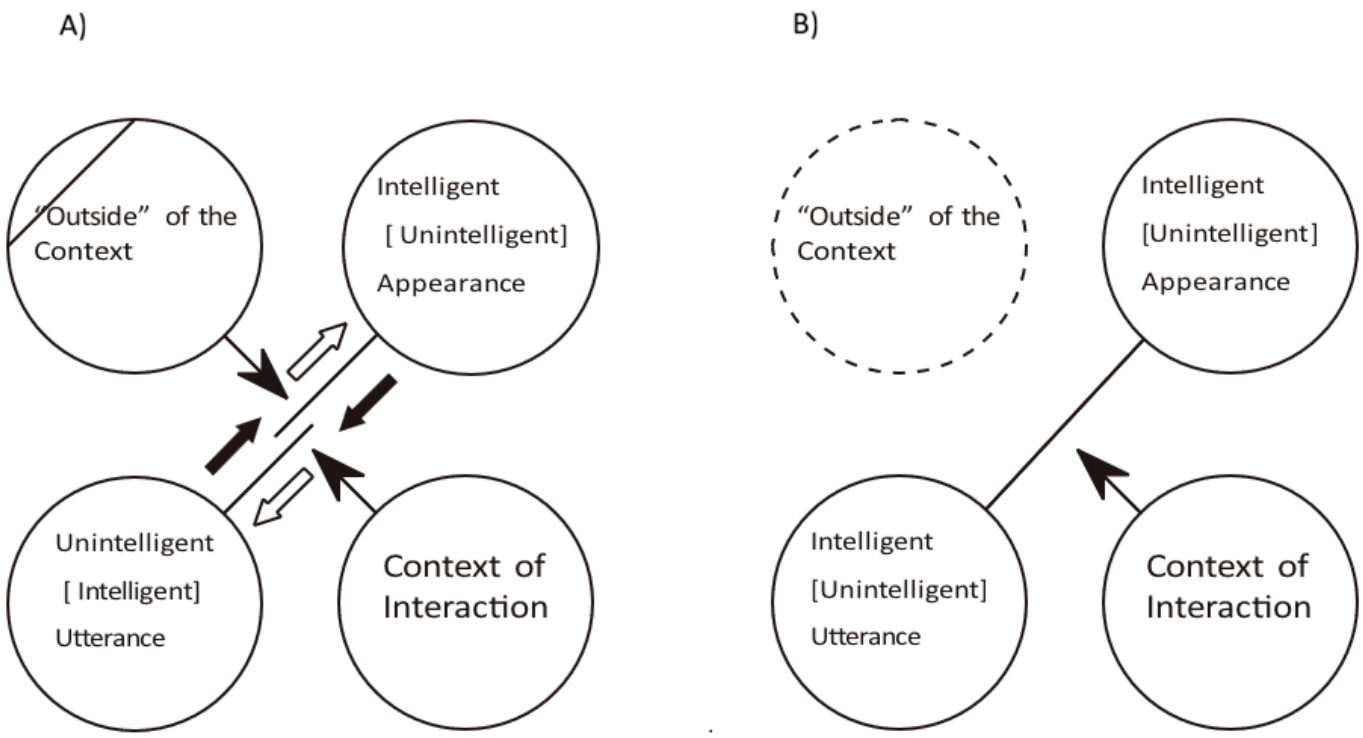

Figure 5

(A) Conceptual diagram of "natural-born intelligence" in this study. The upper right and lower left circles constitute the antinomy. (B) When there is no contradiction between the agent's appearance and speech, antinomy is not constituted, and dynamic changes in the context of interaction do not occur. Created based on Gunji and Nakamura (Gunji and Nakamura 2022) and Gunji and Uragami (Gunji and Uragami 2021), modified to fit this study.