

Why Do We Delegate to Intelligent Virtual Agents? Influencing Factors on Delegation Decisions

Ningyuan Sun

University of Luxembourg
4364 Esch-sur-Alzette, Luxembourg
ningyuan.sun@uni.lu

Jean Botev

University of Luxembourg
4364 Esch-sur-Alzette, Luxembourg
jean.botev@uni.lu

ABSTRACT

Recent intelligent virtual agents (IVAs) are increasingly tasked with critical activities such as financial investment and vehicle control. However, the mechanics behind delegatory behaviors toward IVAs are not fully explored, especially when multiple factors come into play. This paper aims to investigate how different agent-related factors, such as capability and trustworthiness, influence users' decisions on critical-transaction delegation to IVAs. We conducted an experiment constituting a variation of the investment game where participants interact with a robot-like IVA in a virtual environment. We found that, early during the interaction, the informativeness of the agent is more important than other acknowledged factors, such as agent capability. Also, most of these factors have a stable impact on delegation decisions early during the interaction, whereas only a few factors, such as agent usability, have a dynamic impact. Our findings provide guidelines for the design of trustworthy IVA for virtual environments and various levels of immersion.

CCS CONCEPTS

• **Human-centered computing** → **HCI design and evaluation methods; HCI theory, concepts and models; Virtual reality.**

KEYWORDS

intelligent agents, delegation, virtual environments

ACM Reference Format:

Ningyuan Sun and Jean Botev. 2021. Why Do We Delegate to Intelligent Virtual Agents? Influencing Factors on Delegation Decisions. In *Proceedings of the 9th Int'l Conference on Human-Agent Interaction (HAI '21), November 9–11, 2021, Nagoya, Japan*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3472307.3484680>

1 INTRODUCTION

Delegation is omnipresent in human societies. By delegating tasks to agents, principals expect desired expertise (e.g., counseling lawyers, prescribing medicines) or efficiency (e.g., taking taxis). Contemporary agents are predominantly human. However, with the upsurge of artificial intelligence, recent intelligent virtual agents (IVAs) are increasingly capable of accomplishing conventionally human-agent-specific tasks, such as customer service [1], personal

assistance [10], or driving vehicles [3]. Consequently, IVAs are becoming commonplace and replacing human agents.

Many factors have been identified relevant to people's decisions on the delegation to an agent. A large proportion of these factors concern human or virtual agents per se, including, for instance, their adaptivity [4], autonomy [4, 18], capability [12], confidence [5, 23], controllability [9], informativeness [15], informedness [8], politeness [6, 16], predictability [14], representation [22], reputation [5], and visual attractiveness [24]. Other factors are related to the context where delegation occurs, such as the cost of delegation [5] or environment stableness [12], and principals themselves, such as their mood [5], preference [5], or trust [12, 18]. However, it is still unclear how multiple factors in combination shape delegation decisions. This is further complicated by the various contexts in which these factors were identified, including, for example, economic, psychological, sociological, and computer-science scenarios. Without knowing how these factors are organized, inspecting a delegation decision requires examining all the previously mentioned factors. This is hindering future research since considering all the factors is nearly impossible but otherwise would be deemed overlooking.

A structure similar to a neural network (cf. Figure 1) illustrates how different factors influence delegation decisions in combination. The external environment is perceived, processed, and passed to our working memory through our perceptual systems [7]. These pieces of raw information in the working memory constitute the perceptual layer in Figure 1. Based on the raw information perceived, our brains generate various concepts to evaluate the decision on delegation. Each concept comprises a delegation-related factor (illustrated

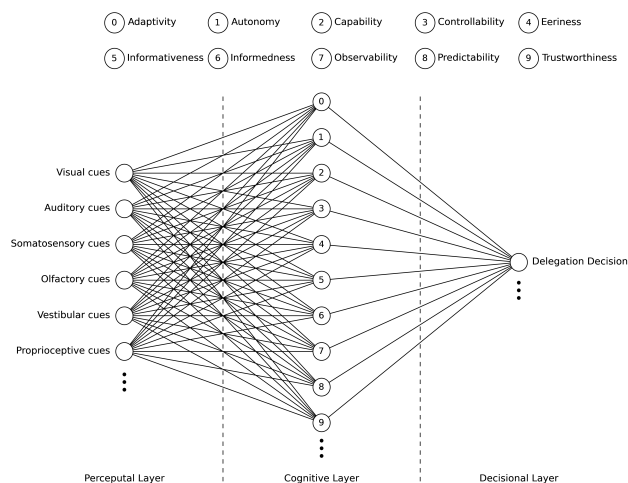


Figure 1: Provisional neural network modeling delegation.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

HAI '21, November 9–11, 2021, Nagoya, Japan

© 2021 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8620-3/21/11.

<https://doi.org/10.1145/3472307.3484680>

as the nodes labeled with digits) and adds up to the cognitive layer. The weight matrix between the perceptual and cognitive layers abstractly represents people’s internal reaction to the raw information perceived. Each individual has a unique weight matrix, modeling that people react differently to the same external stimulus. For instance, iOS users may hold different opinions toward the build-in personal virtual assistant Siri. The cognitive layer is then forwarded to the decisional layer, in which delegation decisions are made. The weight matrix in between denotes the criteria that people rely on to make delegation decisions. For example, while some people may emphasize agent capability, some other people may instead value predictability.

This study focuses on the weight matrix between the cognitive and decisional layers, i.e., the influence of different factors on delegation decisions. We contribute to the topic by attempting to solve a preliminary research question: *which factors are the most relevant for decisions on the critical-transaction delegation to a robot-like IVA?*

2 EXPERIMENT

To answer the research question, we conducted an experiment where the participants interact with an IVA and then take a survey. We will elaborate on the experiment with regard to its design, results, and discussion in this section.

2.1 Design

The experiment, including the interaction and the survey, is realized as a Unity WebGL program and entirely conducted in a browser-based virtual environment to avoid breaks in presence [17]. As shown in Figure 2, the scene is minimally decorated to reduce the potential distraction caused by the environment. The agent has a robot-like visual representation –the robot Kyle [2]– and is dubbed with a synthesized male voice. The agent is animated via motion capture so that the animation appears natural and helps to decrease the level of eeriness [13].

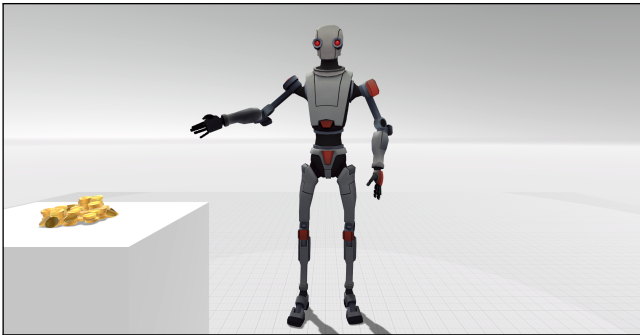


Figure 2: Virtual environment with agent (cf. video [21]).

The interaction with the agent constitutes a variation of the *investment game* to simulate critical transactions. In the virtual environment, the participants are initially endowed with some virtual golden coins. The coins are placed on a virtual table and framed as tokens to exchange for their completion codes after the experiment. The participants can use this code to further exchange for monetary rewards. At the beginning of the experiment, the

agent verbally introduces itself and also the coins’ usage. Following the introduction, the agent makes a proposal to the participants: if the participants gave the agent all the coins they have, the agent would return a different completion code doubling the monetary reward. However, the agent also warns the participants of the risk of failure, in which case the agent returns a completion code worth only half the reward.

After deciding whether to delegate, the participants are informed of an ensuing survey and receive some related instructions. At first, 20 agent-related factors (cf. Table 1) are displayed to the participants in sequence. These factors are considered influential in people’s delegatory behaviors or trusting beliefs. For each factor, the participants are then asked to recall their decision-making process and rank the factor according to the following rules: (1) rank the factor *A* if the factor was consciously, clearly, or explicitly considered when deciding; (2) rank the factor *B* if the factor never came to mind when deciding; (3) rank the factor *C* if the factor is not eligible for any of the two ranks above.

Table 1: Agent-related factors used in the survey

Factor	Opposing Poles	Source
Adaptivity	adaptive – non-adaptive	[4]
Aliveness	animate – inanimate	[19, 22]
Attractiveness	attractive – unattractive	[4, 24]
Autonomy	autonomous – non-autonomous	[4, 18]
Capability	capable – incapable	[12]
Compliance	compliant – non-compliant	*
Confidence	confident – unconfident	[4, 23]
Controllability	controllable – uncontrollable	[12, 18]
Eeriness	eerie – reassuring	[13]
Human Semblance	humanlike – non-humanlike	[20]
Informativeness	informative – uninformative	[4, 15]
Informedness	informed – uninformed	[8]
Moral Quality	benevolent – malevolent	[11]
Observability	observable – unobservable	[12]
Politeness	polite – impolite	[6]
Predictability	predictable – unpredictable	[14]
Reliability	reliable – unreliable	[24]
Reputation	reputable – disreputable	[4]
Trustworthiness	trustworthy – untrustworthy	[12, 18]
Usability	easy-to-use – difficult-to-use	[4]

* Studies on the relationship between IVAs’ compliance and people’s delegation decisions are rare. Still, compliance grows more important with increasing adaptivity.

Once the ranking is finished, all the relevant factors, i.e., the factors that were previously ranked *A* or *C*, are presented to the participants. Since each relevant factor may influence participants’ delegation decisions to a different extent, we then ask them to pick out the *major factors* that largely influenced their delegation decisions and the *minor factors* that influenced their delegation decisions only to a minimal extent. Afterward, the participants are presented with all the relevant factors once again but separately. Each factor is accompanied by a sliding bar. The two endpoints of the sliding bar denote the two poles of a factor as in Table 1. Any point in between is a relative assessment of the factor with

regard to the two poles. The participants are asked to choose a point on the sliding bar representing their assessment for each relevant factor. This allows us to learn the participants' assessment of the agent and force the participants to consciously consider all the relevant factors. Then, the participants have to pick out major factors and minor factors from all the relevant factors once again. The difference before and after the assessment reveals to what extent a short period of conscious consideration can change their impacts on delegation decisions.

2.2 Experiment Results & Discussion

We conducted the experiment with 80 native English speakers recruited from the Prolific crowd platform. The participants are diverse in their age (min = 18, max = 69, mean = 33.3, std = 11.2), gender (46 females, 34 males), countries of residence (mainly Canada, United Kingdom, and United States), occupations, and cultural backgrounds. Most of the participants (70) chose to delegate. Corresponding to the high delegation rate, a large portion (80%) of the factors have received positive assessments on average, while only four factors have received negative assessments, including attractiveness, eeriness, human semblance, and predictability. However, as discussed in Section 1, the current literature struggles to distinguish the essential ones from all the positively assessed factors. The factors' potentially different weights further complicate this. For instance, given the high delegation rate, the median level of the agent's trustworthiness is presumably high since trustworthiness is regarded as one of the most critical factors in delegation. In contrast, the actual level is only slightly higher than the neutral position on the continuum.

To identify the decisive factors, we need to inspect the participants' decision-making process. As shown in Figure 3(a), the participants are diverse in the reported number of factors they have consciously, clearly, or explicitly considered. This number ranges from 2 to 18, and most participants have consciously considered 5 to 12 factors. The wide range indicates that the participants' delegation decisions were made at different levels of awareness. This difference can be caused by, for example, the participants' heterogeneous dispositions of being thoughtful or their various understandings of the term "consciously considering" (from fleeting thoughts to intensive reasoning).

Among all the factors, informativeness, capability, informedness, politeness, reliability, and trustworthiness are the most common factors ranked A (cf. Figure 3(b)). Specifically, a large portion (75%) of the participants have given informativeness rank A, which is significantly higher than other factors such as capability (66%), politeness (58%), and trustworthiness (57%). The prominent awareness of the agent's informativeness is likely due to the participants' desires to familiarize themselves with the virtual environment early during the interaction. Presumably, the participants' attention to the agent's informativeness would decrease as the interaction proceeds. However, the interaction duration is short in this experiment. The participants may only have obtained a limited amount of information and, therefore, are still more concerned with how informative the agent is. The predominance of informativeness is also reflected in its frequency of being selected as a major factor. As shown in Figure 3(c), the agent's informativeness has received the highest number of votes as a major factor, followed by some other common

A-rank factors, such as politeness, and reliability, which also remain major factors after the assessment.

Observation 1 *Decisions on critical-transaction delegation to an IVA are highly relevant to the agent's informativeness early during the interaction, to the extent that the informativeness is more influential in users' delegation decisions than other acknowledged key factors, such as trustworthiness, predictability, etc.*

Figure 3(c) has also revealed the participants' changed attitudes toward several specific aspects of the agent after the assessment. For instance, the agent's usability has been less frequently selected as a major factor and instead has received an increased number of votes as a minor factor. This is further validated by a Wilcoxon signed-rank test ($p = 0.08$, $\alpha = 0.1$). The frequency of the agent's eeriness and reputation being selected as major factors or minor factors has also obviously changed, as Figure 3(c) shows. However, these changes are not statistically supported by the Wilcoxon signed-rank test ($p = 0.31$ for eeriness and $p = 0.11$ for reputation), which may be caused by the small sample size on these two factors.

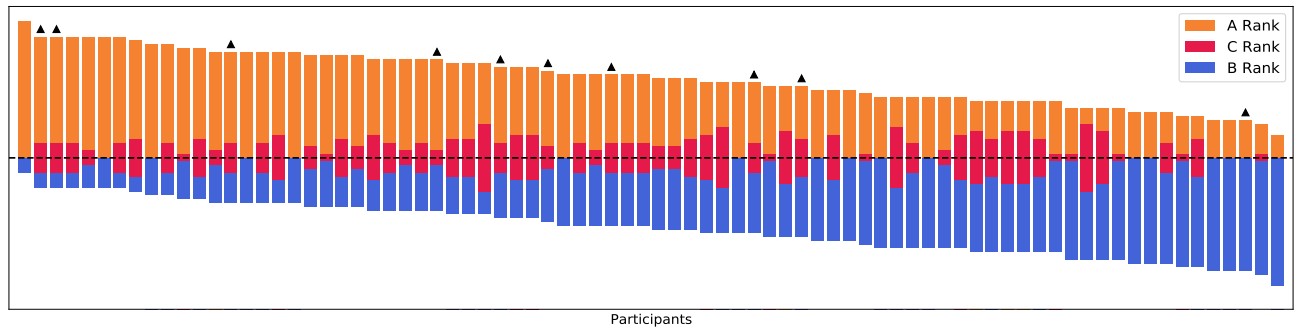
Observation 2 *While the influence of most agent-related factors in delegation decisions is stable early during the interaction, the influence of certain factors, such as usability, eeriness, and reputation, may change rapidly.*

The commonly overlooked factors, i.e., the factors that are ranked B, include moral quality (overlooked by 51% participants), attractiveness (48%), adaptivity (46%), reputation (46%), eeriness (46%), aliveness (46%), and confidence (45%). Among these factors, moral quality, attractiveness, eeriness, reputation, and aliveness are also the most common C-rank factors (39%, 33%, 30%, 28%, and 18%, respectively). The less consideration of moral quality is likely due to the stereotype of machines being user-centered. The agent's visual properties, including attractiveness, eeriness, and aliveness, are demonstrated to influence people's decisions on a subconscious level [24]. The participants were provided with only a limited amount of information about the agent's background or use cases, which may account for their lesser consideration of the agent's reputation. Notably, although these factors are overlooked by approximately half of the participants, there are still many who either consciously considered or were somewhat influenced by them.

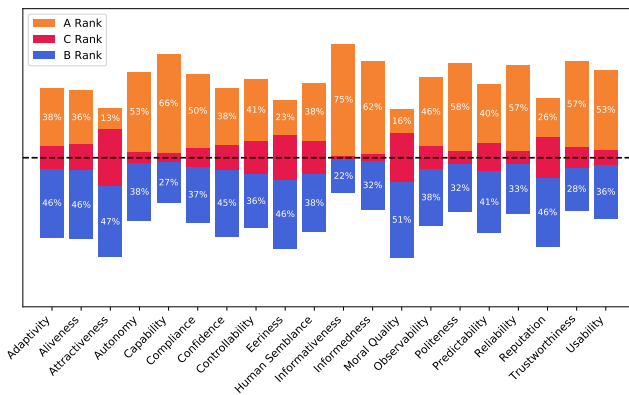
2.3 Limitations

The two observations in Section 2.2 are based on participants' subjective responses, i.e., their consciousness. Although these observations can explain certain delegatory behaviors, they omitted the subconscious influence of agent-related factors on delegation decisions; however, our study provides a baseline to design experiments at a subconscious level. Delegation decisions are also influenced by, for instance, affective and environmental factors. We did not include these factors in our experiment because external factors, such as contexts and users' dispositions, are less controllable than agent-related factors to IVA designers.

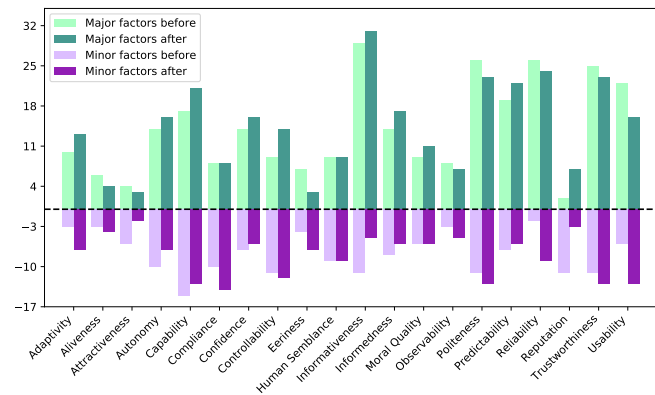
While the contextualization of the experiment is adequate, in reality, decisions on critical-transaction delegation are usually made in more complex situations. Thus, the observations obtained from



(a) Ranking proportions per participant.



(b) Ranking proportions aggregated by factor.



(c) Selection frequency (major/minor, pre/post assessment).

Figure 3: Aggregated experiment results. In Figure 3(a), each bar represents a unique participant. Those who refused to delegate are marked with a triangle on top of the bar. Each bar is composed of a maximum of three colored bars showing the proportion of A-ranked (consciously considered), C-ranked (neither consciously considered nor overlooked), and B-ranked (overlooked) factors. Figure 3(b) follows the same structure except that each bar in Figure 3(b) represents a unique factor and its aggregated ranking results over all participants. Figure 3(c) shows the weighting of the different factors before and after the assessment.

the experiment presented in this paper may only have a limited level of ecological validity and need to be further investigated. When facing decisions on critical transactions to IVAs, people exhibit different levels of willingness to interact with the agent depending on the agent’s form [22]. Other forms of agent representation, such as a zoomorphic agent representation, were not implemented here and need to be investigated in further experiments.

The interaction between the participants and the agent is brief; consequently, the participants cannot accurately evaluate some factors, such as reputation and predictability. These concepts are usually formed over time and, in our case, can only be estimated according to the participants’ experience and their first contact with the agent. Follow-up research should capture the full dynamics of these factors.

Finally, the neural network structure in Figure 1 is only one possible factor model, and not all of them are necessarily considered before making a delegation decision. For instance, most people would not consciously consider all of these factors before making a delegation decision, which is not clearly indicated in the neural network structure.

3 CONCLUSION

This paper presented a survey of the factors that influence human users’ decisions on critical-transaction delegation to IVAs. We experimented with demographically varied participants, exploring the influence of 20 agent-related factors relevant to either people’s trusting beliefs or delegatory behaviors. We investigated the influence of these factors on delegation decisions and derived two significant observations from the experiment results, based on which IVA designers can increase the possibility of an IVA being delegated to by, for instance, supplying information early during the interaction. An IVA’s informativeness is potentially more essential to users’ delegation decisions than its capabilities early during the interaction. While the experiments were conducted in a browser-based virtual environment, our findings also apply to fully immersive settings and affect the design of trustworthy IVAs for VR.

ACKNOWLEDGMENTS

This study has been supported by the Luxembourg National Research Fund (FNR) under grant number 12635165.

REFERENCES

- [1] Martin Adam, Michael Wessel, and Alexander Benlian. 2020. AI-Based Chatbots in Customer Service And Their Effects on User Compliance. *Electronic Markets* 31 (2020), 427–445. Issue 2.
- [2] Kyle Brewer. 2012. *Space Robot Kyle (last accessed: 2021-08-26)*. Unity Technologies. <https://assetstore.unity.com/packages/3d/characters/robots/space-robot-kyle-4696>
- [3] Mark Campbell, Magnus Egerstedt, Jonathan P How, and Richard M Murray. 2010. Autonomous Driving in Urban Environments: Approaches, Lessons and Challenges. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368, 1928 (2010), 4649–4672.
- [4] Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors* 57, 3 (2015), 407–434.
- [5] James M Jenks and John M Kelly. 1986. *Don't Do, Delegate!* Ballantine Books, New York, USA.
- [6] Chris Lam. 2011. Linguistic Politeness in Student-Team Emails: Its Impact on Trust Between Leaders And Members. *IEEE Transactions on Professional Communication* 54, 4 (2011), 360–375.
- [7] Joseph J LaViola Jr, Ernst Kruijff, Ryan P McMahan, Doug Bowman, and Ivan P Poupyrev. 2017. *3D User Interfaces: Theory And Practice*. Addison-Wesley Professional, Boston, USA.
- [8] Carrie R Leana. 1987. Power Relinquishment Versus Power Sharing: Theoretical Clarification And Empirical Comparison of Delegation And Participation. *Journal of Applied Psychology* 72, 2 (1987), 228.
- [9] Brian Lubars and Chenhao Tan. 2019. Ask Not What AI Can Do, But What AI Should Do: Towards A Framework of Task Delegability. In *Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS)*. Curran Associates Inc., New York, USA, 57–67.
- [10] Lorenzo Massai, Paolo Nesi, and Gianni Pantaleo. 2019. PAVAL: A Location-Aware Virtual Personal Assistant for Retrieving Geolocated Points of Interest and Location-Based Services. *Engineering Applications of Artificial Intelligence* 77 (2019), 70–85.
- [11] Roger C Mayer, James H Davis, and David Schoorman. 1995. An Integrative Model of Organizational Trust. *Academy of Management Review* 20, 3 (1995), 709–734.
- [12] Allen E Milewski and Steven H Lewis. 1997. Delegating to Software Agents. *International Journal of Human-Computer Studies* 46, 4 (1997), 485–500.
- [13] Masahiro Mori, Karl F MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine* 19, 2 (2012), 98–100.
- [14] Bonnie M Muir. 1994. Trust in Automation: Part I. Theoretical Issues in the Study of Trust and Human Intervention in Automated Systems. *Ergonomics* 37, 11 (1994), 1905–1922.
- [15] Emrah Onal, James Schaffer, John O'Donovan, Laura Marusich, S Yu Michael, Cleotilde Gonzalez, and Tobias Höllerer. 2014. Decision-Making in Abstract Trust Games: A User Interface Perspective. In *2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*. IEEE, Manhattan, New York, U.S., 21–27.
- [16] Raja Parasuraman and Christopher A Miller. 2004. Trust and Etiquette in High-Criticality Automated Systems. *Commun. ACM* 47, 4 (April 2004), 51–55.
- [17] Susanne Putze, Dmitry Alexandrovsky, Felix Putze, Sebastian Höffner, Jan David Smeddinck, and Rainer Malaka. 2020. Breaking The Experience: Effects of Questionnaires in VR User Studies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–15.
- [18] Nathan Stout, Alan R Dennis, and Taylor M Wells. 2014. The Buck Stops There: The Impact of Perceived Accountability And Control on the Intention to Delegate to Software Agents. *AIS Transactions on Human-Computer Interaction* 6, 1 (2014), 1–15.
- [19] Carolin Straßmann and Nicole C Krämer. 2017. A Categorization of Virtual Agent Appearances and a Qualitative Study on Age-Related User Preferences. In *Proceedings of the 2017 International Conference on Intelligent Virtual Agents*. Springer International Publishing, New York, USA, 413–422.
- [20] Carolin Straßmann and Nicole C Krämer. 2018. A Two-Study Approach to Explore the Effect of User Characteristics on Users' Perception and Evaluation of a Virtual Assistant's Appearance. *Multimodal Technologies and Interaction* 2, 4 (2018), 66.
- [21] Ningyuan Sun. 2021. *Experiment Demonstration Video (last accessed: 2021-08-26)*. University of Luxembourg, VR/AR Lab. <https://youtu.be/CxszlWkh-aE>
- [22] Ningyuan Sun and Jean Botev. 2021. Virtual Agent Representation for Critical Transactions. In *Proceedings of the International Workshop on Immersive Mixed and Virtual Environment Systems (MMVE '21)*. Association for Computing Machinery, New York, NY, USA, 25–29.
- [23] Gunilla A Sundström and Anthony C Salvador. 1994. Cooperative Human-Computer Decision Making: An Experiment And Some Design Implications. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 38. SAGE Publications, Los Angeles, CA, 220–224. Issue 4.
- [24] Beste F Yuksel, Penny Collisson, and Mary Czerwinski. 2017. Brains or Beauty: How to Engender Trust in User-Agent Interactions. *ACM Transactions on Internet Technology (TOIT)* 17, 1 (2017), 1–20.