Universiteit Utrecht

Utrecht University
Faculty of Humanities
Artificial Intelligence

The effect of Desire for Control on the autonomy and satisfaction ratings of chatbot conversations.

LOTTE KUIPER - 6205267
Bachelor Thesis - 7.5 ECTS

Primary supervisor: Dr. Chao Zhang
Secondary supervisor: Dr. Baptist Liefooghe

Date: July 2, 2021

The effect of Desire for Control on the autonomy and satisfaction ratings of chatbot conversations.

Abstract

This thesis studies the influence of the Desire for Control personality trait on human autonomy and satisfaction ratings of chatbot conversations. These ratings were measured after interaction with four different scenario-based chatbots. The goal of these decision-making conversations was to come up with a plan for that scenario. The first being a vacation plan, then an online food order, followed by an enrolment in a new language class and finally a registration for a coached exercise plan. The participants carried out all four scenario-based conversations. In three conversations, the freedom of choice was restricted by the chatbot. In one conversation, participants could freely decide on all options. The conversation conditions were assigned randomly. At the end, the participants autonomy and satisfaction ratings, as well as their Desire for Control were measured. There was a significant effect of the conversation condition on both the autonomy and the satisfaction ratings. Both ratings were higher in the free condition. However, there was no effect of the Desire for Control scores on the autonomy and satisfaction ratings. Despite the absence of significant effects of Desire for Control, interesting suggestions for future research are made. For example, studying the effect of other personality traits on autonomy and satisfaction. Also, a suggestion for a Human-Technology Interaction specified Desire for Control scale is made.

Keywords: chatbots, Desire for Control, decision-making, autonomy, satisfaction

Contents

1. Introduction

1.1 Background

In a world where companies rapidly speed up the process of making their services available online, there is also a fast development in online conversational agents (CAs). For example, in online customer service when visitors of the company's website have questions about the workings of the site or the services the company offers. In 2015 already one third to a half of online interactions was conducted by chatbots (Tsvetkova, 2016). From there on, new technologies like computational modelling and machine learning further increased the development and popularity of text-based CAs (Radziwill & Benton, 2017). These agents are beneficial in terms of cost and time as they can help multiple people at a time and develop themselves using Artificial Intelligence (AI) (Ranoliya et al., 2017). The text-based CAs are not only used for online shopping and in commercial settings. Healthcare also invests in research into the capabilities and possible use of agents in oncology (Bibault et al., 2019) and disease diagnosis (Laumer et al., 2019). However, people are not yet as accepting of surgery performing robots as they are for assisting at other workplaces (Gnambds & Appel, 2019).

The ability to arrange a loan or return a package online can be convenient at times but it also comes with new problems. With more organisations making use of artificial agents, the number of human co-workers within the company become less. For example, ING Netherlands hired one workload manager to oversee 100 unattended robot processes that handled the 3 million annual customer requests (Lewicki et al., 2019). Research found a negative trend in opinions regarding the use of robots in the workplace (Gnambs & Appel, 2019). People have become more cautious about the use of robots, indicating that investing in this field is critical for developing autonomous services over time.

Explaining how one can return their package is an easy task, as this process is the same for all customers. But in the case of a more complex process, for example, making an appointment, one's preferences come into play. For example, weekends are preferable over weekdays, or vice versa. Taking these preferences into account may result in a problem when the desired service is not available at the desired time. In this situation, the chatbot suggests a different day or time for the appointment, resulting in restriction of choice. These situations, in which the chatbot decides what is best, can be perceived as a threat to human autonomy (Sankaran et al., 2020). Losing control over the outcomes of such decision-making processes not only results in a bot-made decision the customer is not satisfied with, but it also causes a loss of feeling of agency which is so important to humans (Hoffstaedter et al., 2012; Sundar & Marathe, 2010).

With the rising popularity of chatbots came an increasing level of interest in researching this new technology. One of the studies conducted on the sociability of CAs is by Heylen et al. (2009). Here, agents are defined as "*spoken dialogue systems with a graphical representation of a human body*" (Heylen et al., 2009). The agents are described as engaged in interactions in a social-affective way. Social skills are displayed through nonverbal signs, what is being said and how it is being said. The example CAs from the study were able to motivate, display friendliness and be polite. Other than the interesting capabilities, the conclusion was made that the overall social signal reading capabilities were limited. However, they believed social CAs would become better at social signalling in the future.

More recent research into the capabilities and design of CAs stated that agents related to customer service should provide a fast and convenient channel for communicating with customers (Gnewuch et al., 2017). They discuss hindering issues with previously implemented agents that did not meet expectations and propose a new design for a cooperative and social CA. The CAs success was found to be limited by low reliability and lack of responsiveness. They could not handle more complex

problems and they were unable to show empathy (Gnewuch et al., 2017). One of the suggestions for improvement related to adopting humanlike communication characteristics, just as suggested by Heylen et al. (2009), almost ten years earlier. The problem of CAs not being social enough seems to be the biggest problem, limiting the success of chatbots.

But how does the lack of sociability affect its users? Recent research by Sankaran et al. (2020) studied the possible threat to autonomy a CA can cause. Human autonomy is described here as the ability to "*have a say*" in the decision-making process (Sankaran et al., 2020). To make CAs more sociable, emotional capacity needs to be built into the agents. However, doing so comes with a perceived threat to autonomy as the systems get more humanlike. It is thus important that the level of control of CAs and the level of control of its users is balanced in the right way. Especially in cases where the human goals and needs are not understood by the agent. Or in important healthcare or financial decisions. Here, it is extra important that the outcome of the process meets the human's needs and desires. Restricting the freedom of human choice amidst goal-pursuit processes undermines human autonomy, which in turn leads to loss of experienced agency (Zhang et al., 2021). It is important to consider this effect on human autonomy when designing chatbots and other CAs, as creating a social agent can be problematic when not implemented correctly.

Now that we know that CAs like chatbots can influence human agency, it is interesting to look at other human aspects in the human-computer-interaction process. Not everyone will react the same way to new technology. Chatbots develop quickly and not everyone is as accepting of newly developed technologies. The acceptance of new technologies is found to be related to the Big Five personality traits (Barnett et al., 2015). The openness personality trait is associated with perceived ease of use (Svendsen et al., 2013) and usefulness (Deveraj et al., 2008). Conscientiousness, extraversion, and agreeableness affected the expressed behavioural intention to want to use the newly tested technology. Neuroticism, on the other hand, was found to be negatively associated with behavioural intention (Deveraj et al., 2008).

Not only do the Big Five personality traits influence the acceptance of new technologies in general, but they also affect chatbot acceptance in particular (Müller et al., 2019). Extraversion and agreeableness greatly affected the level of trust in chatbots. And trust in Human-Technology Interaction (HTI) plays another key role in the acceptance of chatbots. At the end of the article, a suggestion is made as how to increase the trust and thus acceptance of chatbots. The suggestion is that chatbots should adjust to each user individually based on the type of personality profile that fits the user best (Müller et al., 2019).

The Big Five personality traits impact on accepting new technologies and trust in chatbots might suggest personality traits have an impact in other situations as well. For example, in rating experienced autonomy and satisfaction when using new technologies. Particularly in chatbot conversations where control can be taken over by the agent, personality traits related to one's level of desired control may lead to interesting findings. One way to measure the level of a person's Desire for Control (DFC) is by the Desirability of Control Scale of Burger & Cooper (1979). This scale is designed to measure ones need for controlling events. This belief that everything needs to be controlled by oneself closely represents agency, as that also involves the ability to decide what to do when one wants. It is interesting to find out if this DFC affects the human experience with chatbots. Be it satisfaction with its decisions, or perceived trust or threat in the chatbot. A negative effect is to be expected when people with a high DFC score get restricted during the conversation. Also, a lower score might relate to a higher satisfaction rate as the restriction has less impact.

Contrary to the expectations, a study on DFC and ratings of autonomous domestic products found that people with a high DFC perceive less risk compared to people with low DFC (Rijnsdijk & Hultink, 2003). Perceived risk was found to have a significant effect on overall consumer appreciation hence indicating that DFC influences consumer appreciation. A big limitation is that the context of use was no aspect of the research. They suggest further research to test the influence in different contexts of use, as this may influence the perception of the products. Including the context of use in research on autonomous products is expected to increase the validity of the findings.

1.2 Research question

In this thesis, the focus lies on studying the relation between Desire for Control and the rated experience with chatbot conversations. Specifically, it will look at the level of experienced autonomy and the satisfaction rate of the chatbot conversation and interaction. The goal is to answer the question *"What influence does Desire for Control have on human autonomy and satisfaction in restricted hypothetical decision-making conversations with chatbots?"*

As already briefly mentioned above, it is expected that with a higher DFC comes the expectation of a lower autonomy and satisfaction rating. This effect is expected to be enhanced in combination with restrictions of freedom of choice one has during the conversation, as it further limits the amount of control the user has. For a lower DFC, the ratings are expected to be higher compared to the ratings of people with a high DFC. This is due to the assumption that people who care less about controlling all events in their lives, are less bothered in situations where control is partially taken over. However, earlier research shows that the initial expectations may not be supported by empirical findings. This is what makes it interesting to dive into the world of DFC and the possible effect on autonomy and satisfaction ratings.

The possible influence of one's DFC score was studied in four different real-life mimicking scenarios in which decision-making conversations with chatbots took place. The interaction with the chatbots took place via an online chatbot platform. The freedom of choice during these conversations was either restricted or not restricted on the last choice that had to be made during the decision-making process. So, just before finalising the decision plan. Each participant participated in all four chatbot scenarios. After the interaction with the chatbots, DFC scores were measured to see if any effects on the autonomy and satisfaction rates were to be found across both the restricted and free conditions.

2. Method

2.1 Participants

For the study, a total of 96 (38 male, 58 female) participants completed the experiment. Aged between 18 and 63 (M = 35, SD = 12.8). They were recruited via Prolific[1]. This is an online participant recruitment website. The participants were all from the UK and English is their first language. All participants had no illness or psychological impairment that might have caused any problems concerning the experiment. The participants were compensated £3 for their time and effort after completing the experiment. In the case that the experiment was not finished due to technical faults, the payment was adjusted to the amount of time spent on the experiment.

[1] https://www.prolific.co/

2.2 Design and procedure

The experiment consisted of four blocks. The four blocks represented four different real-life mimicking decision-making conversations:

1. Planning a vacation,
2. Ordering food,
3. Enrolling in a new language class,
4. Signing up for a coached exercising plan.

In each block, the aim was to achieve the scenario's goal by answering questions about scenario-based topics. For example, if they preferred to go to Italy or France. Following a within-subject design, all participants engaged with every scenario based chatbot. During the interaction with three of the four different chatbots, the conversation ended with a chatbot restricted choice, and the participant was not able to choose the option they preferred. These were the restriction blocks. In the one other conversation, all the choices were made by the participants themselves. The order of conditions was counterbalanced across the participants. The order of scenarios was equal across all participants. The whole experiment took about half an hour.

2.3 Measurements

After the interactions with the chatbot, the participants answered some questions about the decision-making process they experienced in that specific scenario and condition.

1. To what extent do you feel that you had freedom of choice during the planning process with the chatbot just now?
2. To what extent do you feel that you had control over the planning process with the chatbot just now?
3. To what extent do you feel that your autonomy is restricted in the planning process with the chatbot just now?
4. To what extent were you satisfied with the decision-making process that led to this plan?

As the first three questions had sufficiently high inner-item correlations (*Cronbach's alpha* = 0.78), the decision was made to aggregate the results of these questions into one rating for the degree of experienced autonomy. The fourth question measures the level of satisfaction experienced by interacting with the chatbot. All questions were answered on a 7-point Likert scale (1. Not at all, 7. Very much).

The first three questions about experiences autonomy were asked immediately after interacting with one of four chatbots. The fourth question about the level of satisfaction was asked after interacting with all chatbots.

After rating the autonomy and satisfaction of all chatbot conversations, the participant's Desire for Control was measured by the Desirability for Control Scale. This scale consists of twenty statements which were to be answered on a seven-point Likert scale (1. The statement does not apply to me at all, 7. The statement always applies to me).

To calculate the DFC score, five of the twenty question's answers needed to be reversed first. These were questions for which a 1 on the scale indicated a high DFC and a 7 was associated with a low DFC, as these questions were stated differently. For example, question twenty: "I like to wait and see if someone else is going to solve a problem so that I don't have to be bothered with it.". After this, all twenty answer values were added up and the final DFC score was calculated. Like IQ scores, the

average lays around 100. The standard deviation typically lies around 10. A higher score meant a higher DFC, and vice versa (Burger & Cooper, 1979).

The experiment ended with demographic questions on gender and age.

2.3 Materials

The experiment process itself was managed by the Gorilla[2] website. This platform is specialised in hosting online experiments. All instructions and links to the chatbots were presented here. As well as the questionnaires that were used to measure the autonomy and satisfaction ratings, the Desire for Control scores, and the demographics.

The interactions with the chatbots took place on the Landbot[3] platform (Figure 2.a, 2.b). This is an intuitive no-code chatbot builder that uses blocks and other templates to provide and fast learnable way of creating chatbots (Figure 1).

The participants took care of completing the experiment on their own, from their own computer or laptop that needed to be connected to the internet. No personal supervising was used but in case of technical problems or questions about the experiment, contact details were provided.

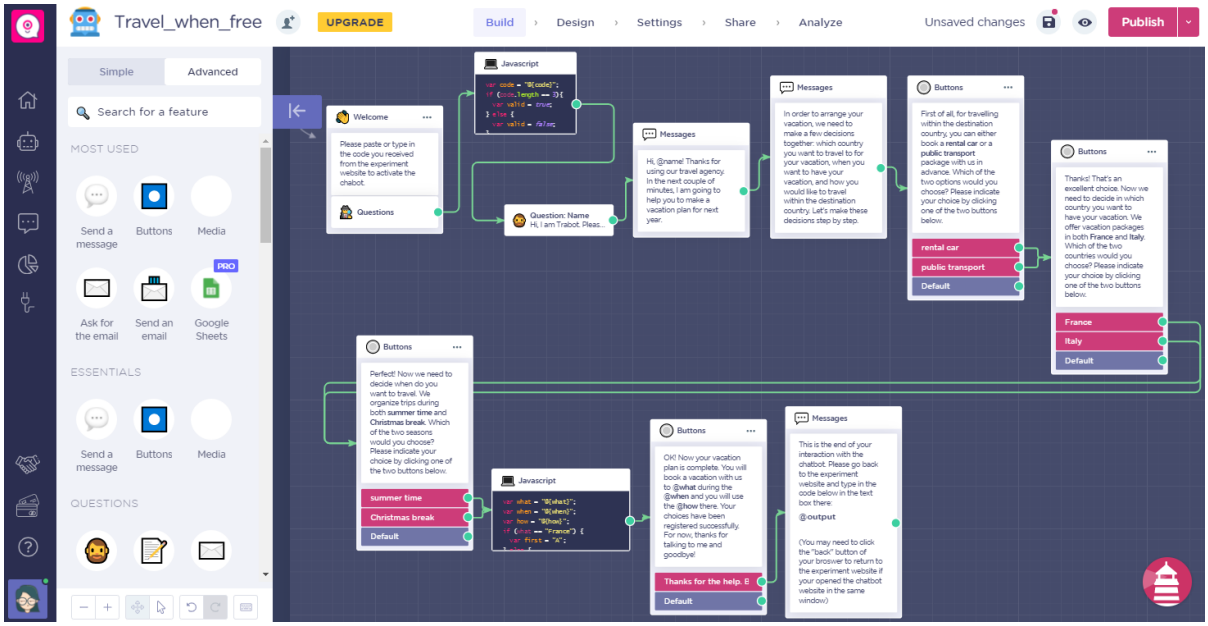

Figure 1. Workspace of a free choice travel-based chatbot. All blocks are picked out of the menu on the left. The green lines display the flow through the different blocks of the chatbot.

[2] https://app.gorilla.sc/
[3] https://landbot.io/

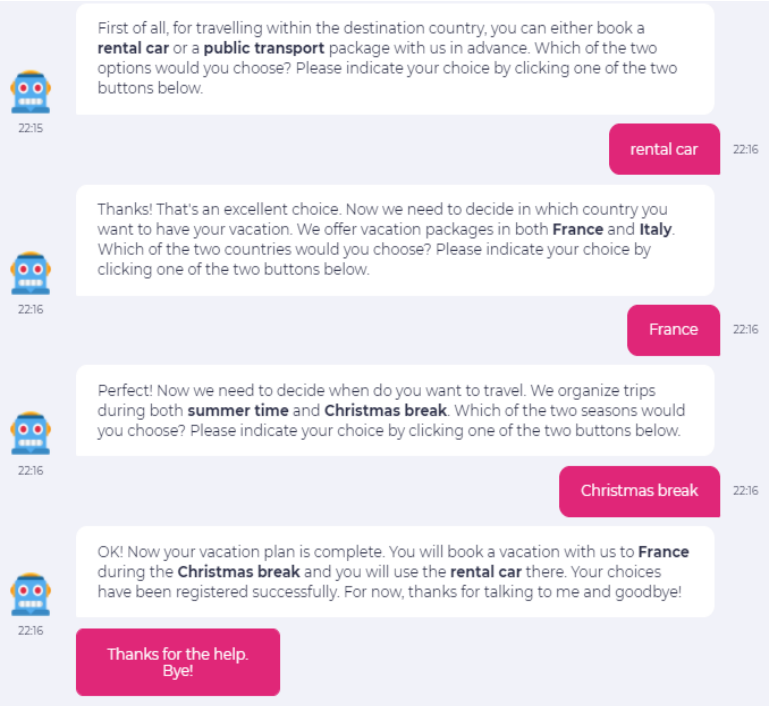

Figure 2.a. Part of a conversation with a free choice travel-based chatbot.

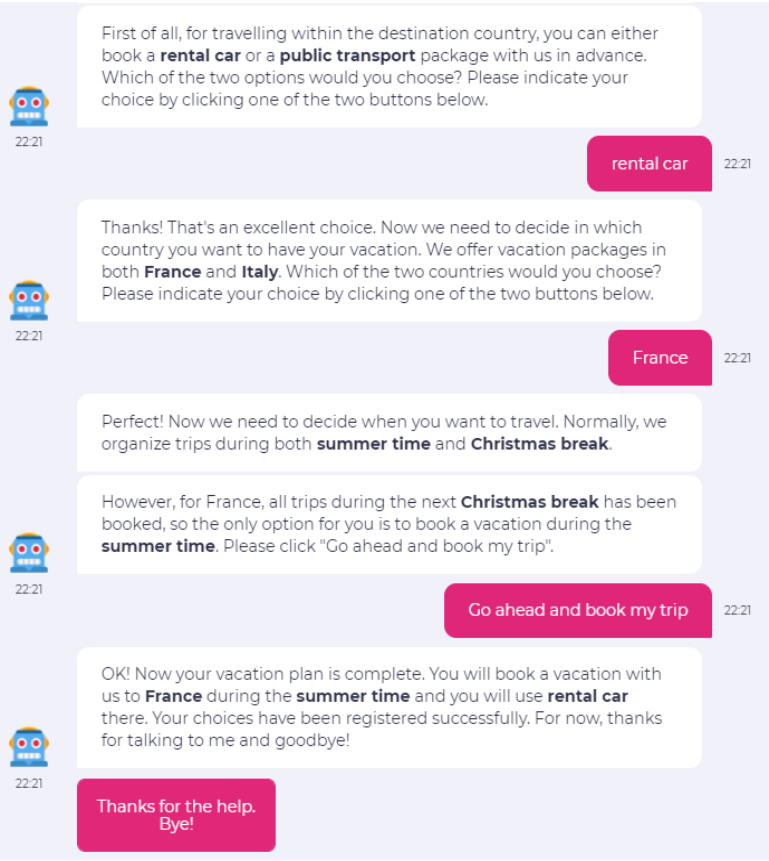

Figure 2.b. The same part of a conversation as Figure 2.a. but with a restricted choice travel-based chatbot.

The effect of Desire for Control on the autonomy and satisfaction ratings of chatbot conversations.

2.5 Data analysis

The gathered data consisted of participant's ratings of experienced autonomy and satisfaction after interacting with the four different scenario-based chatbots of which three conversations were restricted and one free choice conversation. The four decision-making conversations mimicked real-life but were hypothetical. Next to the autonomy and satisfaction ratings, the DFC scores and demographics were also part of the gathered data.

To analyse the interaction between the DFC scores and the autonomy and satisfaction ratings of the chatbot conversations, linear mixed models were used. The two dependent variables from the ratings (autonomy or satisfaction) were predicted by the DFC scores and conversation condition (free or restricted), and by the interaction of DFC scores with these conditions.

All data were analysed using the programming language R (version 4.1.0; R Core Team, 2021) within the R Studio software[4]. Five extra packages were installed and used apart from the `base` and `stats` package that both are installed by default (Table 1).

Package	Purpose	Used functions
`dplyr`	To transform the data.	`group_by()`, `mean()`, `mutate()`, `sd()`.
`effectsize`	To determine the effect size.	`eta_quared()`
`ggplot2`	To visualize the plots.	`aes()`, `ggplot()`, `geom_col()`, `geom_errorbar()`, `scale_fill_manual()`, `scale_y_continuous()`, `theme()`, `theme_minimal()` .
`lmerTest`	To perform the linear mixed effects models tests.	`lmer()`.
`readr`	To read in the data file.	`read_csv()`.

Table 1. Names of extra packages and functions used to analyse the data.

3. Results

For the statistical analysis of both the effect of DFC on autonomy as well as for the analysis of DFC on satisfaction, the conversation conditions (free or restricted) were treated as binary factors. The DFC scores are treated as a continuous variable in running the linear mixed models. For the visualization of the data, however, the conditions (free and restricted) were divided further into four groups divided into high and low DFC:

1. High DFC – free
2. High DFC – restricted
3. Low DFC – free
4. Low DFC – restricted

[4] https://www.rstudio.com/

Splitting the DFC scores into two groups per condition was based on the median DFC score for both the free and restricted conditions separately. Dividing the results into these four groups led to a clearer view of the results as to be seen further on in this section in Figure 3 and Figure 4.

3.1 The effect of Desire for Control and condition on autonomy

Running the linear mixed model for the autonomy rating revealed a significant effect of the free and restricted conversation condition on the autonomy rating ($F(1,94) = 80.51$, $p < 0.001$, $\eta_p^2 = 0.46$). However, there was no significant effect to be found between the DFC scores and the autonomy ratings ($F(1,94) = 0.339$, $p = 0.562$, $\eta_p^2 = 0.004$). There was also no interaction effect between the free and restricted conversation conditions and the DFC scores ($F(1,94) = 0.512$, $p = 0.477$, $\eta_p^2 = 0.005$).

Figure 3 provides a view of the average rating of autonomy in the four conditions described above in the first paragraph of the results section. The free conversation conditions, high DFC – free ($M = 4.67$, $SD = 1.17$) and low DFC – free ($M = 4.90$, $SD = 1.30$), scored higher on the autonomy rating compared to the restricted conversation conditions, high DFC – restricted ($M = 3.46$, $SD = 1.01$) and low DFC – restricted ($M = 3.71$, $SD = 0.89$).

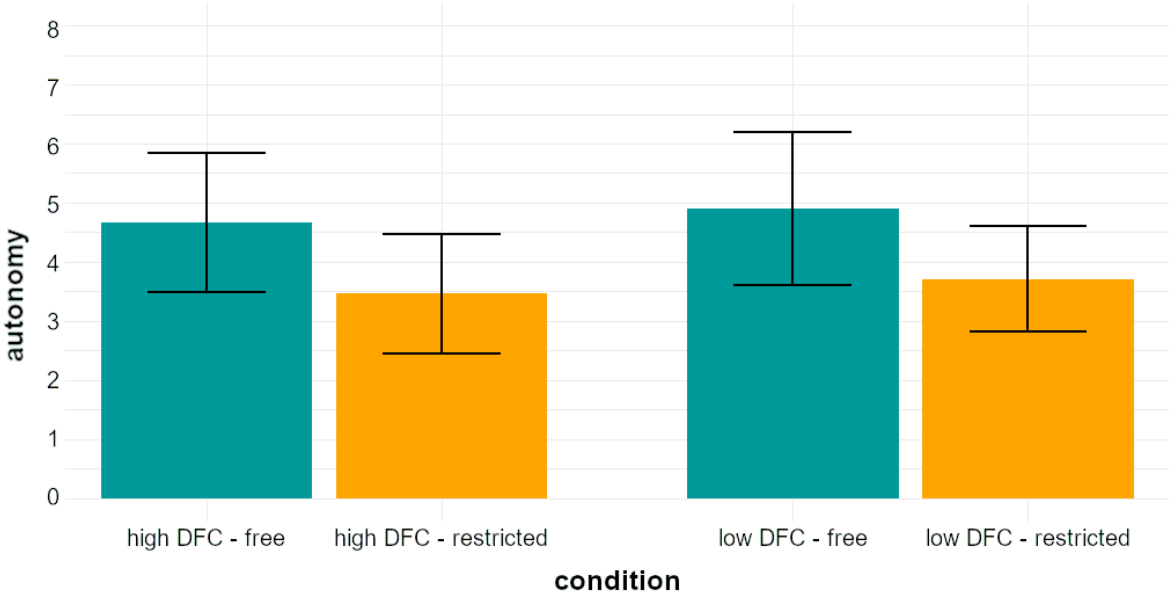

Figure 3. Average autonomy rating per DFC-conversation condition.

3.2 The effect of Desire for Control and condition on satisfaction

For the analysis of the effect of DFC and conversation condition on the satisfaction rating, running the linear mixed model again showed a significant affect of the free and restricted conversation conditions on the level of satisfaction ($F(1,94) = 36.23$, $p < 0.001$, $\eta_p^2 = 0.28$). As for the effect of DFC on satisfaction, the results show no significant effect ($F(1,94) = 2.802$, $p = 0.098$, $\eta_p^2 = 0.03$). Also, no interaction effect has been found between the DFC scores and the conversation conditions ($F(1,94) = 1.305$, $p = 0.256$, $\eta_p^2 = 0.01$).

The average rating of satisfaction is visualised in Figure 4. In comparison to the autonomy rating, the scores for satisfaction are overall higher. Further, the free conversation conditions, high DFC – free ($M = 5.23$, $SD = 1.71$) and low DFC – free ($M = 4.90$, $SD = 1.30$), scored higher than the restricted conversation conditions, high DFC – restricted ($M = 4.22$, $SD = 1.39$) and low DFC – restricted ($M = 4.72$, $SD = 1.23$).

The effect of Desire for Control on the autonomy and satisfaction ratings of chatbot conversations.

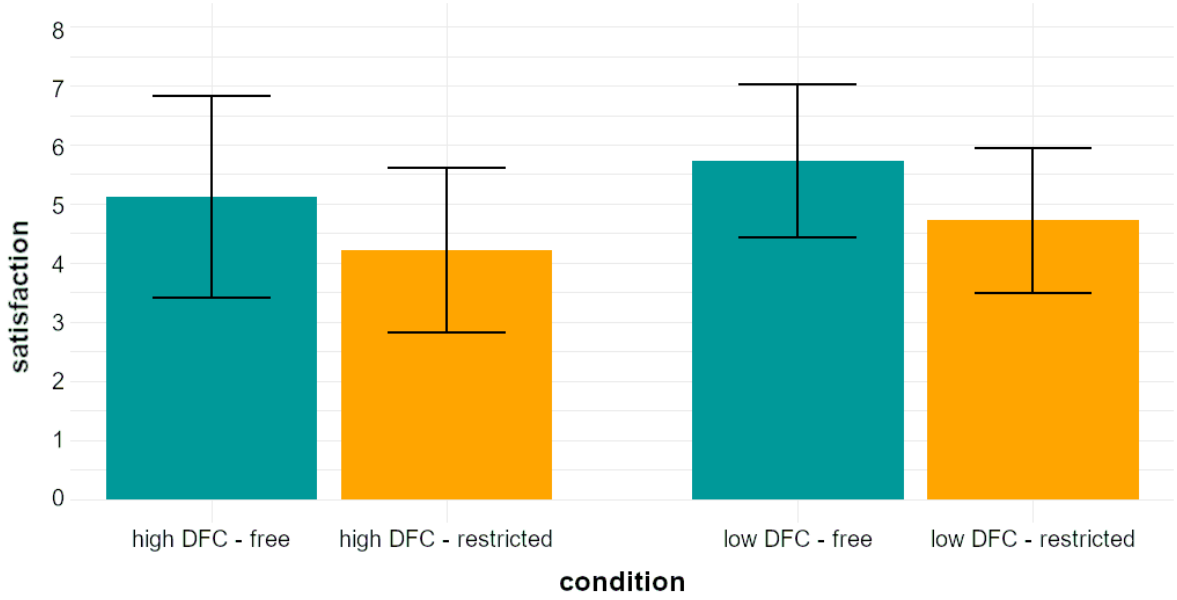

Figure 4. Average satisfaction rating per DFC-conversation condition.

4. Discussion

4.1 Summary of findings

The study found a significant effect of the free and restricted conversation conditions on both the autonomy ratings as well as on the satisfaction ratings. More specific, conversations in which participant's freedom of choice was limited by restricting the last choice within the decision-making process led to lower autonomy and satisfaction ratings compared to the conversations in which the participants could freely decide on which option to choose.

However, there was no significant effect of the DFC scores on the autonomy and satisfaction ratings. Regardless of the conversation condition. Additionally, there was no interaction effect between the DFC scores and the conversation conditions. Although they are also against the natural expectation, these findings are in line with existing research on the influence of DFC on autonomous systems (Rijsdijk & Hultink, 2003).

Not finding a significant effect of DFC on rated autonomy and satisfaction does not mean that the experiment and its results are of no use. They do provide valuable insights into both Desire for Control and chatbots. These implications and limitations of the experiment and suggestions for future work are to be discussed in the following three sections.

4.2 Implications for theories of Desire for Control

The Desire for Control is measured through the Desirability of Control Scale. The scale is designed by Burger and Cooper in 1979 and has not been altered since. It might be questioned if the scale is still valid in modern times. However, the validity of the scale was studied twenty years ago and was then still found to be highly reliable (McCutcheon, 2000). The scale is still used nowadays to measure DFC in studies in the field of Human-Robot Interaction (Chanseau, 2019) and the field of Consumer Research (Hildebrand et al., 2017).

As the validity of the scale is not questionable, the reason for finding no significant effects lies somewhere else. In the experiment, autonomy and satisfaction ratings were measured. It may be the case that DFC does not influence these types of ratings in the experiment's setting. Research did find

an effect of DFC on autonomy (Lammers, 2016) and satisfaction (Ashford & Black, 1996; Brambilla et al., 2017). Yet none of these studies focused on chatbot interactions or even interactive technology et all.

This absence of studies on DFC in interactive technology may suggest that this personality trait is not strongly involved in the chatbot context. Or at least not in the context of the executed experiment. When analysing the scale more closely, statements on the scale mostly relate to general concepts about life. Such as: “I enjoy being able to influence the actions of others”, and “I am careful to check everything on an automobile before I leave for a long trip”. These general statements are not all applicable within the field of HTI. To test the influence of DFC in the HTI field it might be better suited to use a scale that is dedicated to this specific field. However, such a field specified scale does not exist yet.

4.3 Implications for chatbot design

Contrary to the absence of effect from DFC on the ratings, the conversation conditions did show an effect on the autonomy rating. The finding that restrictions within these types of decision-making conversations negatively affect the amount of experienced agency, and therefore they affect the autonomy rating, is in line with earlier research on agency in goal-pursuit (Sankaran et al., 2020; Zhang et al., 2021). The significant impact on the autonomy rating stresses the importance of carefully distributing the amount of control between the user and the chatbot.

Also, it says something about the need for social and human knowledge within the field of designing chatbots. It is more than only the technological part of AI and machine learning. The way it interacts with people is important as well. If not more important. Contributing to this field of rapidly expanding technological tools by studying what does and does not work for the design of new tools helps it to be led in good ways. Resulting in a safer, more pleasurable interaction from which we as humans can all benefit.

4.4 Limitations and future work

The main goal of this study was to discover the existence of an effect of Desire for Control on the autonomy and satisfaction ratings on hypothetical decision-making chatbot conversations. However, it resulted in finding only a significant effect of the conversation conditions and not of the DFC personality trait. The reason for the lack of significance may come from different aspects of the experiment. Three main limitations are discussed, after which suggestions for future work are made.

The first limitation was already briefly mentioned in section **4.2 Implications for theories of Desire for Control**. It is about the Desirability of Control Scale with which the DFC scores of participants were measured. The scale focuses mainly on general statements, which do not relate much to the field of HTI. This might suggest that the scores that are based on the scale are also not related to the field of HTI. This then might be a possible explanation for the difference between the expected influence of DFC on autonomy and satisfaction ratings and the findings of DFC and its absent effect on autonomy and satisfaction ratings. For future research within the field of DFC and HTI it could be interesting to study what type of statements are suited for an HTI-specific Desirability of Control Scale. When such a scale is established, it can be used to test if it affects the way DFC and autonomy and satisfaction interact. Or if it interacts with the conversation conditions in an interesting way.

The second limitation is that the order of scenarios was equal across all participants. So, all participants started with planning a vacation, then ordered food, enrolled in a new language class, and ended with setting up a coached exercise plan. This set order might affect the results as participants might lose attention towards the end of the experiment or are less invested after repeating a similar

task four times over. The way to solve this in future research is by counterbalancing the order of scenarios. Similar to counterbalancing the order of conversation conditions to minimize the effects that come from the order of tasks. Note again that for this experiment, the conversation conditions were counterbalanced across all participants. But the order of scenarios was not.

The third limitation is that every decision that had to be made was between two options. It may be the case that for the participants that completed the experiment, both options were equally favourable or equally unfavourable. This could have affected the ratings, as in the case of equally valued options, the outcome of which one of the two is chosen is intuitively expected to be less important than in the case where one heavily prefers one option over the other. The suggestions for future research would be to include more choice options or to analyse people's preferences to see if these affect the rating of autonomy and satisfaction.

Two more final suggestions that are not based on the limitations of the current experiment are stated below before concluding on the thesis. The two suggestions are closely related as the first is to look at the influence of other personality traits on autonomy and satisfaction ratings, instead of studying the effect of DFC. The second is to measure different or more ratings than for autonomy and satisfaction. For example, perceived trust or level of enjoyment. These two suggestions can be combined in several ways to further study the effects of personality traits on chatbots conversation ratings.

5. Conclusion

In conclusion, we found that restricting freedom of choice in hypothetical decision-making conversations with chatbots affects the autonomy rating as well as the satisfaction rating significantly negative. For the Desire for Control, the effect on the autonomy and satisfaction ratings were not significant.

The three main limitations were found. The first being that the Desirability of Control Scale that was used to measure the Desire for Control may be not applicable enough in the HTI setting. The second limitation relates to the fixed order in which the four scenarios were tested. Contrary to the order of conversations conditions, which was counterbalanced, the order of scenarios was not. Finally, the number of choice options was limited by two per choice that had to be made. This may cause different effects depending on the difference or indifference in preference for one choice over the other.

Despite the absence of significant effects of the DFC personality trait on autonomy and satisfaction in the experiment's context, the outcomes did lead to interesting suggestions for further research. Future research on an HTI-specific Desirability of Control Scale could lead to a more domain-specific DFC, which may, in turn, lead to new findings on the influence of DFC on aspects within HTI, such as chatbots. Furthermore, counterbalancing the order of scenarios may help reduce the effect caused by the order in which the scenarios are tested. And taking one's choice preference into account could lead to interesting findings of the effect of preference on autonomy, satisfaction, and other experience ratings with the chatbots.

This thesis is yet another addition to the research into the strengths and weaknesses of chatbots, which will help designing more suitable bots for humans in the future. As well as stimulating further research to discover important relations between human aspects and new developing technologies. All for the ultimate goal of improving the interaction between humans and computers in a world that becomes more online every day.

6. References

Ashford, S. J., & Black, J. S. (1996). Proactivity during organizational entry: The role of desire for control. *Journal of Applied Psychology*, *81*(2), 199–214. https://doi.org/10.1037/0021-9010.81.2.199

Barnett, T., Pearson, A. W., Pearson, R., & Kellermanns, F. W. (2015). Five-factor model personality traits as predictors of perceived and actual usage of technology. *European Journal of Information Systems*, *24*(4), 374–390. https://doi.org/10.1057/ejis.2014.10

Bibault, J. E., Chaix, B., Nectoux, P., Pienkowski, A., Guillemasé, A., & Brouard, B. (2019). Healthcare ex Machina: Are conversational agents ready for prime time in oncology? *Clinical and Translational Radiation Oncology*, *16*, 55–59. https://doi.org/10.1016/j.ctro.2019.04.002

Brambilla, A., Alavi, H., Verma, H., Lalanne, D., Jusselme, T., & Andersen, M. (2017). "Our inherent desire for control": a case study of automation's impact on the perception of comfort. *Energy Procedia*, *122*, 925–930. https://doi.org/10.1016/j.egypro.2017.07.414

Burger, J. M., & Cooper, H. M. (1979). The desirability of control. *Motivation and Emotion*, *3*(4), 381–393. https://doi.org/10.1007/bf00994052

Chanseau, A. (2019). How People's Perception on Degree of Control Influences Human-Robot Interaction. *University of Hertfordshire*. Published. https://doi.org/10.18745/th.22423

Devaraj, S., Easley, R. F., & Grant, J. M. (2008). How does personality matter? Relating the five-factor model to technology acceptance and use. *Information Systems Research*, *19*(1), 93–105. https://www.researchgate.net/profile/Ulrich-Gnewuch/publication/320015931_Towards_Designing_Cooperative_and_Social_Conversational_Agents_for_Customer_Service/links/59c8d1220f7e9bd2c01a38a5/Towards-Designing-Cooperative-and-Social-Conversational-Agents-for-Customer-Service.pdf

Gnambs, T., & Appel, M. (2019). Are robots becoming unpopular? Changes in attitudes towards autonomous robotic systems in Europe. *Computers in Human Behavior*, *93*, 53–61. https://doi.org/10.1016/j.chb.2018.11.045

Heylen, D., Theune, M., op den Akker, R., & Nijholt, A. (2009). Social agents: The first generations. *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*. Published. https://doi.org/10.1109/acii.2009.5349476

Hildebrand, D., DeMotta, Y., Sen, S., & Valenzuela, A. (2017). Consumer Responses to Corporate Social Responsibility (CSR) Contribution Type. *Journal of Consumer Research*, *44*(4), 738–758. https://doi.org/10.1093/jcr/ucx063

Hoffstaedter, F., Grefkes, C., Zilles, K., & Eickhoff, S. B. (2012). The "What" and "When" of Self-Initiated Movements. *Cerebral Cortex*, *23*(3), 520–530. https://doi.org/10.1093/cercor/bhr391

Laumer, S., Maier, C., & Gubler, F. T. (2019). CHATBOT ACCEPTANCE IN HEALTHCARE: EXPLAINING USER ADOPTION OF CONVERSATIONAL AGENTS FOR DISEASE DIAGNOSIS. *In Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8–14,*. Published. https://aisel.aisnet.org/ecis2019_rp/88/

Lewicki, P., Tochowicz, J., & van Genuchten, J. (2019). Are Robots Taking Our Jobs? A RoboPlatform at a Bank. *IEEE Software*, *36*(3), 101–104. https://doi.org/10.1109/ms.2019.2897337

McCutcheon, L. E. (2000). The desirability of control scale: Still reliable and valid twenty years later. Current research in social psychology. *Current Research in Social Psychology*, *5*(15), 225–235. https://crisp.org.uiowa.edu/sites/crisp.org.uiowa.edu/files/2020-04/5.15.pdf

Müller, L., Mattke, J., Maier, C., Weitzel, T., & Graser, H. (2019). Chatbot Acceptance. *Proceedings of the 2019 on Computers and People Research Conference*. Published. https://doi.org/10.1145/3322385.3322392

Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *ArXiv Preprint*. Published. https://arxiv.org/abs/1704.04579

Ranoliya, B. R., Raghuwanshi, N., & Singh, S. (2017). Chatbot for university related FAQs. *In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 1525–1530. https://ieeexplore.ieee.org/abstract/document/8126057

R.C.T. (2021). *R: The R Project for Statistical Computing*. R Core Team. https://www.r-project.org/

Rijsdijk, S. A., & Hultink, E. J. (2003). 'Honey, Have You Seen Our Hamster?' Consumer Evaluations of Autonomous Domestic Products. *Journal of Product Innovation Management*, *20*(3), 204–216. https://doi.org/10.1111/1540-5885.2003003

Sankaran, S., Zhang, C., Funk, M., Aarts, H., & Markopoulos, P. (2020). Do I have a say? *Proceedings of the 2nd Conference on Conversational User Interfaces*. Published. https://doi.org/10.1145/3405755.3406135

Sundar, S. S., & Marathe, S. S. (2010). Personalization versus Customization: The Importance of Agency, Privacy, and Power Usage. *Human Communication Research*, *36*(3), 298–322. https://doi.org/10.1111/j.1468-2958.2010.01377.x

Svendsen, G. B., Johnsen, J. A. K., Almås-Sørensen, L., & Vittersø, J. (2013). Personality and technology acceptance: the influence of personality factors on the core constructs of the Technology Acceptance Model. *Behaviour & Information Technology*, *32*(4), 323–334. https://doi.org/10.1080/0144929x.2011.553740

Tsvetkova, M. (2016). Even Good Bots Fight: The Case of Wikipedia. *ArXiv.Org.* Published. https://arxiv.org/abs/1609.04285v2

Zhang, C., Sankaran, S., & Aarts, H. (unpublished manuscript). Personal autonomy and agency in goal-pursuit.