MASTER

The Effect of Choice Characteristics on the Delegation of Subjective Decisions to AI

van der Weide, Fabiènne P.

Award date:
2023

The Effect of Choice Characteristics on the Delegation of Subjective Decisions to AI

Fabiènne Pascalle van der Weide

1004980

Eindhoven University of Technology

In partial fulfilment of the requirements for the degree of

Master of Science

in Human Technology Interaction

Supervisors:

1st supervisor: dr. C. Zhang

2nd supervisor: dr. ir. M.C. Willemsen

3rd assessor: prof. dr. C.C.P. Snijders

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Chao Zhang. Not only did he guide me through every step of this project, but he did so while being on paternity leave during a big part of it. I am grateful for all the insightful feedback I got while designing, conducting, and analysing the experiment. I want to thank him for encouraging and challenging me, helping me to accomplish the best result possible.

I would also like to thank Martijn Willemsen for making the time in his schedule to be my second supervisor. His expertise and critical evaluation helped a lot in the design and final writing phase of the process and gave me new inspiration to think about certain elements on another or deeper level.

I am very thankful to my friends and family for their love and support during this process. A special thanks to family and friends who participated in my numerous pre-tests. I could not have done it without them.

Finally, I would like to thank all of the participants in my study for their time and valuable contributions. This work would not have been possible without their willingness to sacrifice a moment of their time.

Abstract

Every day, people have to make tons of decisions. As humans often experience various difficulties when making decisions, delegating these to artificial intelligence (AI) can offer a solution. While earlier research mainly focussed on delegation to other humans or on how decision difficulty influences delegation, van den Bergen (2023) did research into the effect of choice characteristics on delegation to AI. In the current study, we replicated her work while making a few adjustments to address some open issues. In an online experiment (N=80), delegation rates for six different conditions were investigated (e.g. one option dominates the other, two options are equally attractive, or a large trade-off has to be made) in two contexts: deciding between two hotels and deciding between different numbers of coins in a gamble game. The results showed that delegation rates were lowest when one alternative is dominant compared to the other, followed by a trade-off condition in which one option has more balanced values and the other more extreme values. Delegation rates were higher for conditions with two options that are equally good or equal but neither good nor bad, as well as for trade-offs between two options that display a mirror-image of each other in terms of attribute values. Delegation rates were highest for decisions between two options that are equally bad. Based on additional qualitative analyses into people's motivation to delegate to AI, we recommend taking individual differences between users into account when applying automated decision-making. Moreover, more research must be done into what influences trust in AI and how to increase it.

Contents

1. Introduction

Imagine that you are planning a city trip to Paris. When looking for a hotel, you find one very expensive hotel, but really close to the city centre, while on the other hand, you find an option that is a lot cheaper, but also more than an hour away from the city centre. How would you choose between these two options, when neither one of them dominates the other and each of them is better than the other in one dimension?

Every day, people make tons of decisions that are comparable to the situation described above. We have to make trade-offs when choosing an apartment to live in, when comparing jobs, when buying a new laptop, even when choosing a mode of transport or when doing groceries. Humans are often unable to make optimal and efficient decisions, as they can be indecisive or experience conflicts (e.g. Luce et al., 1999; Rassin, 2007). Having too many options to choose from can even make us less happy than having a small, limited number of options, according to choice-overload hypothesis (e.g. Iyengar & Lepper, 2000; Scheibehenne et al., 2010).

Technology can be a useful aid in decision-making situations, as it does not experience the same difficulties as humans do when selecting one from multiple options. With the rapid development of artificial intelligence (AI), tools like ChatGPT are becoming more and more popular (Brockman et al., 2023). Next to exploiting AI as an aid in various situations, one could also take it one step further and fully delegate decisions to AI. This is also called automatic decision-making (ADM), which includes aiding human decision-makers by algorithmic data processing to fully automated decision-making across a wide variety of contexts and situations. (Araujo et al., 2020; Davenport & Harris, 2005; Newell & Marabelli, 2015). These situations include but are not limited to social contexts like hiring and firing employees, robotic diagnostics in health care, and recommender systems in various contexts (e.g. Braun, 2019).

ADM is more efficient than human decision-making, as AI does not experience the same struggles as humans do when subjected to difficult choices. For this reason, it would be beneficial to employ ADM in many more daily situations. However, if we want to make optimal use of ADM, it is important to know if and when people are willing and when they are not willing to use these kinds of systems. As there exist different types of difficult decisions, it is plausible to assume that the willingness to delegate is not equal for these different situations. While some literature exists on people's willingness to delegate to other humans (e.g. Steffel & Williams, 2018) and the comparison between delegation to humans or AI (e.g. Candrian & Scherer, 2022; Leyer & Schneider, 2020), an examination of the potential influences of conflict type on delegation to AI is still limited.

The current study investigates the effect of choice characteristics on the delegation of subjective choices to AI. Examples of investigated choice characteristics are whether a dominant alternative is present or not, whether options are positive or negative compared to a neutral reference point, and what the size of the trade-off is. The following section presents a literary base of the hypotheses and methods presented further in this thesis. An explanation of the different experimental contexts, stimuli design, and statistical tests will be given, after which the results will be presented and discussed. Finally, we will describe some limitations, theoretical and practical implications, and advice for future research.

2. Related work

2.1 Decision-making and decision difficulty

In all choices, some kind of trade-off is involved; You need to choose one thing at the expense of another (Shaddy et al., 2021). At the same time, not all decisions are equally difficult. Overall, decisions including a large trade-off are more difficult than decisions including a clear dominant alternative (Shafir et al., 1993). According to Bettman et al. (1993), this is caused by increased cognitive difficulty. A well-known theory regarding decision-making is utility theory (Fishburn, 1968): people tend to choose the option with the highest expected utility. This utility is calculated using a value function, which is subjective and can be different for each individual. In conclusion, when one option is dominant, its utility is higher, and this option is likely to be chosen over other options with a lower utility.

When the utilities of two options are very close, decisions become more difficult. Still, people are often capable of applying strategies to decide between options. According to prospect theory, outcomes of decisions are expected to be coded as either a gain or a loss relative to a neutral reference point, in which losses loom larger than gains (Kahneman & Tversky, 1979; Tversky & Kahneman, 1991). Adding to this, Chatterjee and Heath (1996) found that decision difficulty ratings are lower for trade-offs on positive attributes (i.e. approach-approach conflicts) than for trade-offs on negative attributes (i.e. avoidance-avoidance conflicts).

If people regard two options as equal, it is called perceived similarity. People may base this perceived similarity on the differences between values of different attributes or on the commonalities between options and on how well the options fulfil their goals (Xu et al., 2013). There are different methods to measure preference and perceived similarity. Two popular methods are matching and choice. With matching, participants are asked to fill in a value for which two options are equal. With choice, participants are asked to choose the option that is favoured over the other. The two methods, however, show different results and preference orderings, which can be explained by assuming different weighing of the attributes. In choice, the prominence effect is present (the most important attribute receives more weight), which is not the case with matching (Tversky et al., 1988; Willemsen, 2002). Important to note when applying the matching method in a study, is that people are reluctant to match on values that have a negative valence compared to the values of the other option for this attribute (Willemsen & Keren, 2002).

Next to establishing that decisions between options with comparable utilities are more difficult than when a dominant alternative is present, larger trade-offs also appear to be more difficult than small trade-offs or decisions between two equally attractive options (Chatterjee & Heath, 1996; Kim et al., 2013). As two similar options can be regarded as means for reaching the game goal, people are likely to be indifferent between these two options (Xu et al., 2013). Interesting to note is that perceived similarity increases when slight differences within an attribute (e.g. price) are present, compared to both options having the exact same price (Kim et al., 2013).

Scholten & Sherman (2006) proposed a model which is aimed at describing the relationship between choice difficulty and trade-off size. This double-mediation model has an inverted U-shape, meaning that choice difficulty is smallest for very small or very large trade-offs, but larger for trade-offs of moderate size. The reason for this might be that larger trade-offs are more difficult because of the greater sacrifices that need to be made when favouring one option over the other, but with even larger trade-offs, justifying a decision to one self is easier than with slightly smaller trade-offs. Comparable effects are also found by Kim et al. (2013), who found that the perceived similarity of options decreases choice difficulty and choice avoidance. At the same time, Willemsen et al. (2016) found that greater diversity between several options decreases choice difficulty and enhances choice satisfaction.

An example of how people resolve conflicts in trade-offs can be found in the prominence effect (as briefly mentioned earlier). This effect explains how the more important (prominent) attribute within a choice receives more weight in a decision and that people tend to choose the less negative value on the attribute with a negative valance (Willemsen & Keren, 2002). Other phenomena that could be occurring when making decisions and influence the choice that people will make are, for instance, loss aversion (Kahneman & Tversky, 1979; Tversky & Kahneman, 1991), risk aversion (Tversky & Kahneman, 1981) and extremeness aversion (Simonson & Tversky, 1992).

2.2 Deferring and delegating decisions

When faced with a difficult decision, one strategy is to defer the choice. The decision to defer a choice is influenced by the absolute difference in attractiveness among alternatives provided. A no-choice option is preferred if new alternatives are added with equal overall attractiveness. The preference for a no-choice option decreases when adding inferior alternatives (Dhar, 1997). More difficult trade-offs are deferred more than an easy choice (with a dominant alternative) (Dhar & Simonson, 2003).

When not deferring a decision completely, one could also delegate a decision. Delegation might reduce indecisiveness and choice difficulty when choosing between a continuously rising number of purchase alternatives (Broniarczyk & Griffin, 2014). People like to delegate decisions to another person to avoid feelings of regret (Steffel & Williams, 2018). More difficult decisions are delegated more often than easy ones; the importance of the decision itself does not matter. The reason for delegation is merely to avoid responsibility, not the possibility of a poor outcome. On the other hand, other researchers found that managers are less likely to delegate decisions to others when the stakes are higher, compared to less important decisions (Leana, 1986; Yukl & Fu, 1999).

Next to delegation to other people, one can also delegate a decision to AI. People experience AI as less self-interested than a human, which means there are equal delegation rates for losses and gains (while one might be less willing to delegate to another human when the decision includes losses) (Candrian & Scherer, 2022). This can be explained using the control premium, which states that people are more likely to make their own decisions if their aim is maximizing pay-off, driven by a need for control (e.g. Bobadilla-Suarez et al., 2017; Bohnet & Zeckhauser, 2004; Butler & Miller, 2018; Candrian & Scherer, 2022; Owens et al., 2014).

Other studies comparing delegation to humans versus computers are not always consistent (Candrian & Scherer, 2022; Dawes et al., 1989; Dietvorst et al., 2015, 2018; Leyer & Schneider, 2020; J. Logg, 2018; J. M. Logg et al., 2019; Longoni et al., 2019, 2020; Pezzo & Beckstead, 2020; Schmidt et al., 2020; von Eschenbach, 2021). People have mixed opinions about the fairness and usefulness of automated decision-making and different studies find both algorithm aversion and algorithm appreciation in different domains. Overall, delegation to another person seems to be preferred for subjective contexts or human tasks and algorithms for objective contexts or mechanical tasks (Castelo et al., 2019; Lee, 2018; J. Logg, 2018). Algorithm aversion also depends on personal factors, such as age, technology familiarity, or extraversion (Mahmud et al., 2022).

2.3 Delegation and choice characteristics

Combining knowledge about decision-making and delegation (to AI), van den Bergen (2023) was, to our knowledge, the first to examine the effect of choice characteristics on the delegation of subjective choices to AI. She investigated how people's average delegation rates to AI differed across five different decision conditions with distinct choice characteristics. Specifically, she looked at the difference in delegation rates between an easy decision (with one dominant option) compared to different types of difficult decisions (two options are equally bad or equally good, or a large trade-off has to be made in which one option has attribute values that are close together or the attribute values of both options are far apart). The hypotheses were tested in two different contexts; one that entailed the choice of a hotel (which is a personal, hypothetical scenario) and one that entailed gambling decisions in a coin game (which had real monetary consequences).

Delegation rates were measured by presenting participants with 15 choice tasks for both contexts. In the hotel context, participants had to choose between two hotels based on two attributes. The attribute pairs were the price of the room versus time travelling to the city centre, the cleanness of the room versus how quiet the hotel and surroundings are, and the quality of the breakfast versus the climate control in the room. In the coin context, participants were asked to choose between a certain number of blue and red coins, after which one of the two values would be doubled with a chance of 50%. For all choice tasks, they could choose between hotel/option A or B, or choose to delegate the decision to AI (which, they were told, was trained based on their personal preferences).

Using a logistic regression model, the difference in delegation between the decision conditions was investigated for both the hotel and the coin context separately. Van den Bergen (2023) found that, in the hotel context, delegation rates were lower for the dominant condition, compared to equally good, equally bad, and both trade-off conditions. In the coin context, delegation rates were lowest in the dominant and the first trade-off condition (in which there is one balanced option), followed by the equally good and equally bad conditions, and highest in the second trade-off condition (with all extreme values).

The current research is a close replication of van den Bergen (2023), with a few small adjustments. The previous study gives a good base of what patterns to expect in terms of delegation rates across various types of decision conflicts, but some questions remain unanswered and additional questions have arisen. For instance, what is the reason that there are bigger differences in delegation rates between the decision conditions in the coin context compared to the hotel context? Can the results of the coin context be generalized when another rule for the value of the coins is applied? And will differences in delegation rates between decision conditions be larger when the differences between stimuli values are also more extreme?

3. Research question and hypotheses

To summarize, we are replicating the study of van den Bergen (2023) while making several changes to address some open issues. The main goal of these changes is to make the two contexts more comparable in terms of how the values of the stimuli are presented. We wanted to do this to check if the differences found by van den Bergen (2023) are caused by the difference in the setting of the decision or by the difference in how the stimuli values are interpreted.

3.1 Changes replication study

First of all, we tried to make the decision conditions (dominant, equally good/neutral/bad, trade-offs 1 and 2) more distinguishable from each other and more comparable across contexts. Van den Bergen (2023) calculated the expected utility of all choice options and controlled for this across conditions. This means that she made sure that the utility of all stimuli was approximately the same; The utility of all conditions was above a neutral reference point and that of equally bad was below this reference point. With our study, we wanted to have all conditions around a neutral reference point instead of above it. For this reason, a new choice condition is added; next to equally good and equally bad, we also have an equally neutral condition. This new condition also means that equally good is above the neutral reference point and equally bad is below it. This also allowed us to investigate the differences between gains and losses in more depth than van den Bergen (2023) was able to do.

Designing all stimuli in such a way that they are around a neutral reference point also means that we had to change the stimuli values in the coin condition, making the value 0 this neutral reference point. Previously, all stimuli in the coin context were positive, except in the equally bad condition (van den Bergen, 2023). We changed it so that the average utility is always 0 for all stimuli, meaning that every trial has both positive and negative values, meaning a chance to either gain or lose coins (equally good has only positive values and equally bad had only negative values, the average utility across these two conditions is 0 again).

Second, we were interested in whether making the differences between conditions bigger results in a bigger difference in delegation rates, too. This is why, in the hotel context, the differences between the different conditions were made bigger and more comparable to the coin context in terms of the presentation of the values. This means, for example, that the values in the trade-off condition with all extremes are more representative as a mirror image of each other, rather than just a very large trade-off between two attributes. The dominant alternative in the dominant condition is made even more dominant and the values in the equally good and bad conditions are made more extreme. In the coin context, we made the values of the attributes in the second trade-off condition (where one option is the mirror image of the other option) slightly different from each other, rather than being the exact mirror image. The goal was to make the stimuli more realistic and more similar to the slightly dissimilar values in the hotel stimuli, as well as that it increased the perceived similarity of the options (Kim et al., 2013).

Third, since the pair with cleanness and noise levels of the hotel context turned out to not be able to represent equal choices well, this pair was changed to cleanness versus bed comfort. Van den Bergen (2023) did not find balanced choice distributions for the stimuli in her cleanness versus noise level pair. Participants seemed to value a cleaner hotel over a more quiet environment in most cases, regardless of the stimuli values and decision conditions. We used a new attribute to replace the noise level attribute to overcome this issue.

Fourth, the ranges of the attribute values in the hotel context were changed to make them more realistic or intuitively interpretable. The price attribute was changed to €50-€600 instead of €45-€300 to make them a more realistic presentation of reality and allow for more extreme values. The range for breakfast, climate control, cleanness, and bed comfort was changed to 0-10 instead of 0-7 to make the values more intuitively interpretable and to make the values more similar to the coin context (which has values ranging from -10 to 10). The range for travelling time to the city centre remained 0-80 minutes.

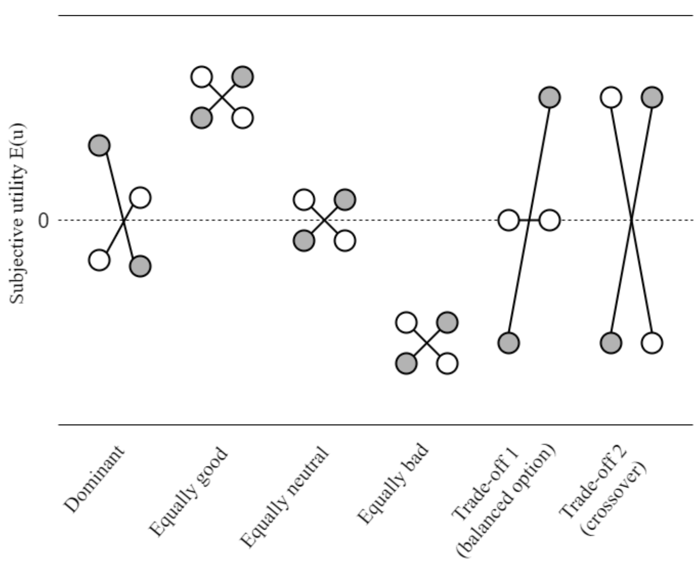
Finally, the rule for what happens to the values of the coins after a decision is made in the coin context was changed back to the rule from the original paradigm (Gillan et al., 2015). This was done because we wanted to investigate if results generalize if we applied a slightly different game rule. The rule used to be that, after a decision was made, one coin (either red or blue, with a 50 per cent chance) doubled in value while the other would remain the same. Now, the rule is that either the red or the blue coin becomes worthless instead of being doubled in value (still keeping the value of the other coin unchanged).

3.2 Research questions

As briefly mentioned in the introduction, the goal of this research is to investigate the difference in delegation to AI for six decision situations that are qualitatively different from each other (see Figure 1). The research question, based on van den Bergen (2023), is in which decision situations (e.g. one option dominates the other, both are equally bad) people are more willing to delegate decisions to AI. The previous study found different patterns in delegation rate between five different decision situations across the two different contexts. The current research will dive even more in-depth into the question of how delegation is context-dependent by making minor adjustments, trying to replicate the basic results and ruling out some alternative hypotheses. An additional research question of the present research, compared to the previous one, is how the context of the decision (hotel versus coin) and the presentation of information influence delegation to AI.

Figure 1

Abstract Visualization of Different Decision Conditions and Subjective Utility in the Coin Context



Note: The circles represent the attribute values of a choice. Within each condition, circles of the same colour (connected by a line) represent one choice option. The same conditions are the same in the hotel context, but the neutral average reference utility is not 0, but 5.1 (normalized to a scale of 0-10 for attributes for which this is not the case (i.e. price and time)).

3.3 Hypotheses

In line with the findings of previous research, we expected participants to delegate difficult decisions more often than easy decisions including a clear dominant alternative (Dhar & Simonson, 2003; Shafir et al., 1993; Steffel & Williams, 2018; van den Bergen, 2023).

H1. People delegate less often to AI in situations where one choice option dominates the other, compared to situations where the two options are equal (or close to equal) in terms of their subjective utilities.

Van den Bergen (2023) found lower delegation rates for the trade-off condition with a balanced option compared to the equal conditions, but higher rates in the other trade-off condition (with extreme values for both options), in the coin context. As, in the current study, differences between the hotel and coin context are expected to be smaller, we expect the same pattern for both conditions.

H2. People delegate less often to AI in situations where a large trade-off has to be made in which there is one option with extreme values for both attributes and one balanced option, compared to situations when two equal options only involve a small trade-off.

H3. People delegate more often to AI in situations where a large trade-off has to be made in which one option is the "mirror image" of the other option, compared to situations when two equal options only involve a small trade-off.

Furthermore, we do not expect differences between the equal conditions (equally good, equally neutral, equally bad), in line with van den Bergen (2023).

H4. When two options are equal (or close to equal) in terms of their subjective utilities, the valence of the two options does not influence the likelihood of delegation to AI (i.e. there are no differences between equally good, bad, and neutral conditions).

Put differently, we expect the following pattern in delegation rates for the six conditions: large trade-off cross-over > equally good = equally bad = equally neutral > (large trade-off with one balanced and one extreme option, dominant). We don't have a specific prediction about the difference between the last two conditions.

4. Method

4.1 Design

The study is a one-factor (choice condition) within-subject design. The dependent variable is whether people delegate a decision to AI or not. The six choice conditions, as explained in the previous section, are (1) dominant, (2) equally neutral, (3) equally good (4) equally bad, (5) a trade-off between one balanced and one extreme option, and (6) a trade-off between two options that are each other's mirror-image. We tested the hypotheses in two different contexts. One involves a hypothetical scenario regarding the choice of a hotel and the other involves a gamble game involving real monetary consequences. The study was approved by the institutional ethical review board (HTI Ethical Review Board – experiment ID 1809).

4.2 Participants

80 participants (47 female; 29 male; 1 other (i.e. non-binary); 3 not specified; M_{age}=28.55; SD_{age}=14.47; *range* = 18-84 years old) were recruited from the TU/e JFS Participant Database (randomly selected) and complemented with convenience sampling through personal networks. The original goal was to get 100 participants (see Appendix A, p.75), but this was not possible due to time constraints, which is why we stopped data collection at 80 participants. None of the participants had vision problems or did not have a sufficient command of the English language. The participants filled in information regarding their occupation (15 full-time job; 31 part-time job; 34 no job), their income (45 less than €1.000 per month; 19 €1.000 -

€2.000 per month; 8 €2.000 - €3.000 per month; 5 €3.000 – €4.000 per month; 3 more than €4.000 per month), and whether they are a student or not (56 student; 24 not student). Participants gave consent (in an online form) and received monetary compensation for their participation.

4.3 Stimuli

For both contexts, the used stimuli were based on the stimuli used by van den Bergen (2023), to which the changes as described in the previous section were applied. In the first[1] context, participants were asked to choose between two hotels that differed in terms of two different characteristics. These attributes are presented in pairs, which are price and time travelling to the city centre, cleanness and bed comfort, and breakfast and climate control. Van den Bergen (2023) decided to select specifically these attributes because they are inherently uncorrelated. We only changed the second pair from cleanness versus noise levels to cleanness versus bed comfort, as van den Bergen (2023) found that results in her study were not reliable for this pair. In the coin context, different numbers of red and blue coins (ranged -10 to 10) were presented and people had to select the option (A or B) which they favoured over the other. After each choice was made, one of the coins (either blue or red) would be devalued to be worth nothing with a 50 per cent probability for both kinds of coins.

In a number of pre-tests, exchange rates were calculated for each of the hotel attribute pairs. Based on these tests, the following exchange rates were assumed: 3.5 for price divided by time, and 1 for cleanness divided by bed comfort as well as for breakfast divided by climate

[1] During the experiment, the order in which the contexts are presented is randomized

control rating. In the coin context, the assumed exchange rate between red and blue coins was always 1, as there should not be a subjective difference between the two.

In the design of the stimuli values, we controlled for these assumed exchange rates as well as expected utility. Literature suggests that taking the square root as the utility function is suitable for the calculation of subjective attributes (Galanter, 1962), which is why van den Bergen (2023) decided to use this utility function. This means that for the hotel context, the utility function is $E(u) = \sqrt{x_1} + \sqrt{x_2}$ (after standardizing the price and time attribute values to a scale 0-10 to match the other attributes) and for the coin context, the utility function is $E(u) = 0.5 \cdot \sqrt{x_1} + 0.5 \cdot \sqrt{x_2}$.

All stimuli were designed in such a way that the expected utility was approximately equal for the dominant condition (when taking the average of the two options in a trial), the equally neutral condition, and both trade-off conditions. The equally good condition trials had a higher expected utility and the equally bad condition a lower one. Averaged, the equally good and equally bad conditions combined have the same expected utility as all other conditions. Controlling for the average expected utility was done to ensure that found effects between conditions are indeed caused by the different decision conditions and not by the difference in utility. For the hotel context, the average expected utility for the stimuli is 5.1. For the coin context, the average utility is 0, meaning that for all conditions there are both positive and negative values (in other words, a chance to either gain coins or lose coins) except for the equally good (only positive values) and equally bad (only negative values) conditions.

All stimuli were tested and validated in a series of pre-test, in which we looked at the number of people that chose either option A or B in all trials. If a clear preference for either one of the options became apparent in any of the choice conditions except the dominant condition, the values were re-evaluated and changed accordingly. In the same way, if the dominant alternative in the dominant conditions appeared to not be favoured by the majority of participants, values were changed for the next pre-test. For an overview of all final stimuli, see Table 1 (p. 34) to Table 4 (p. 39).

4.4 Experimental tasks

The online experiment was programmed in lab.js, an online tool for building studies (Henninger et al., 2022). The experiment consists of several tasks within the two different choice contexts, including a point allocation task, various matching tasks, decision-making tasks, and open questions regarding decision strategy and motivation to delegate. Within the hotel context, participants took part in the point allocation task, several matching tasks, and choice tasks. Within the coin context, participants took part in various matching and choice tasks again, but not in a point allocation task.

4.4.1 Point allocation task

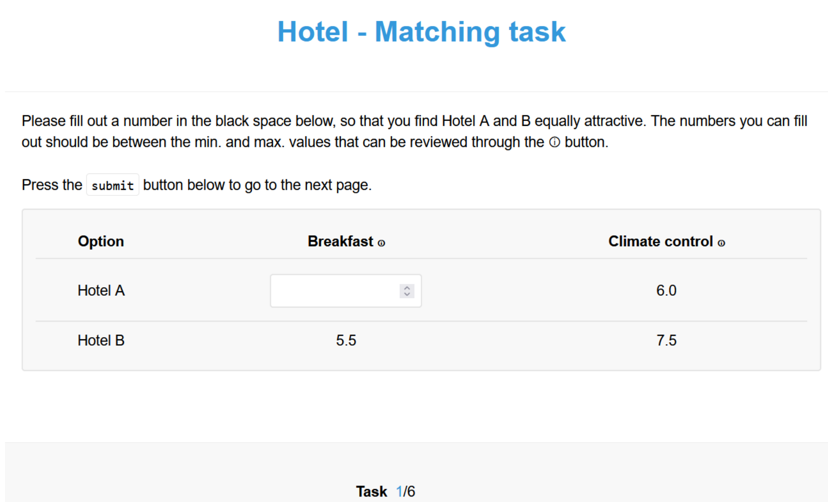
As mentioned before, the attributes on which the choices in the hotel contexts are based are the price of the room (50-600 euros), time to the city centre (0-80 minutes), cleanness rating (0-10), bed comfort rating (0-10), breakfast rating (0-10), and climate control rating (0-10). In the point-allocation task, participants were asked to divide 100 points over these attributes, indicating how important they found each of the different characteristics. This task is included to verify the exchange rates of the three attribute pairs used in the choice tasks, but more importantly to ensure that participants believe the AI is being trained based on their preferences. In reality, there is no AI and choices made by the “AI” are pre-programmed (the dominant option in the dominant condition and random in all other conditions).

4.4.2 Matching tasks

A second type of task included to ensure a successful "real AI" manipulation is the matching task. Here, participants were presented with an example of a choice task in which one of the attribute values is left empty. Participants had to fill in a value for which both options (or hotels) were equally attractive to them (as shown in Figure 2 and Figure 3). As people are reluctant to match on values that have a negative valence compared to the values of the other option for this attribute (Willemsen & Keren, 2002), we ensured to always let them fill in the more positive value in the tasks. In total, six matching tasks were included in both contexts of which the order was randomized for each participant. The matched-on attribute (either A or B) was counterbalanced within participants.

Figure 2

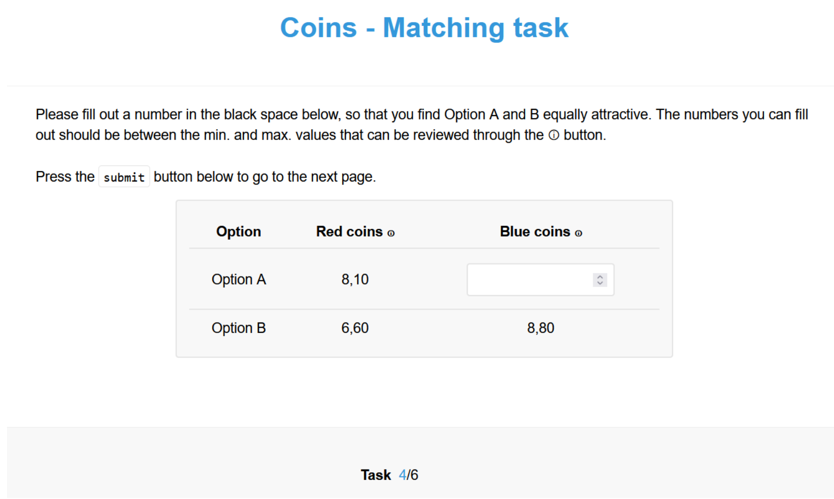
Example of a Matching Task in the Hotel Context



Note: Here, participants had to fill in a breakfast rating for Hotel A, for which they found Hotel A and Hotel B equally attractive.

Figure 3

Example of a Matching Task in the Coin Context



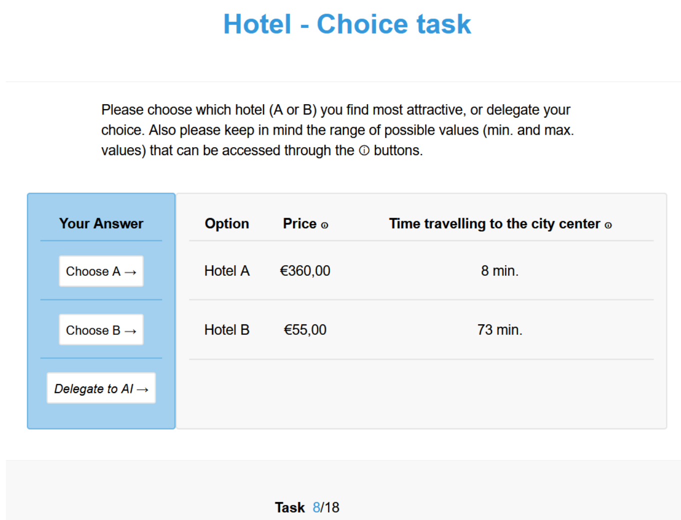
Note: Here, participants had to fill in how many blue coins for Option A would make Option A and Option B equally attractive.

4.4.3 Choice tasks

In the choice tasks, participants filled in their preference for either option A or B or they chose to delegate the decision to AI (see Figure 4, Figure 5). In both contexts, 3 choice tasks per decision condition were included (18 tasks per context, 36 in total). After each decision, participants could immediately see the outcome of their (either delegated or not) choice.

Figure 4

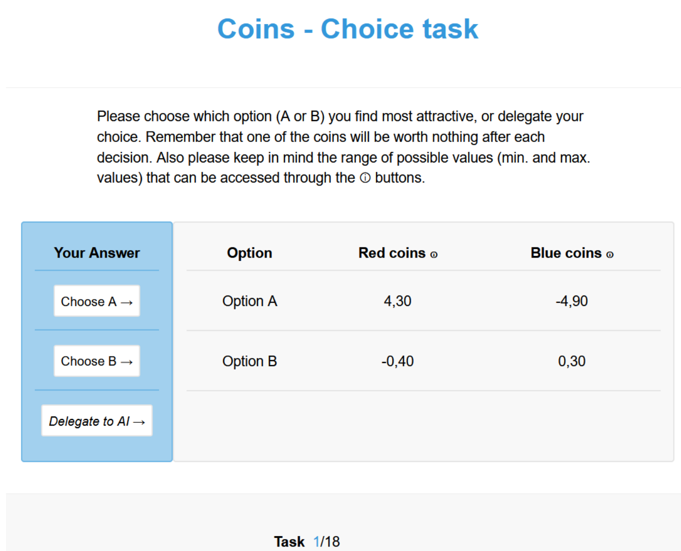
Example of a Choice Task in the Hotel Context



Note: Here, participants had to either choose between hotels A and B or decide to delegate the decision to AI. This is an example from the trade-off 2 condition.

Figure 5

Example of a Choice Task in the Coin Context



Note: Here, participants had to either choose between option A and B or decide to delegate the decision to AI. This is an example from the trade-off 1 condition.

4.4.4 Open questions

After the matching and choice tasks of both trials, a couple of open questions regarding people's personal decision strategies and motivation to delegate a decision to AI were asked. For both contexts, four trials from the choice tasks were shown to the participant again; one from the equally good condition, one from the equally bad condition, and one from both trade-off conditions. The dominant condition was not included, because it is less interesting to investigate how people make a decision in this situation (since most people will automatically go for the dominant alternative). The equally neutral condition was not included either, as the assumption was made that it would not present more additional insights than when only investigating the similarities and differences between equally good and equally bad and comparing those to the trade-off conditions.

4.5 Measurements

Only for the hotel context, participants filled in importance scores for each of the six attributes. In both choice contexts, we measured delegation rates to AI for the specified choice conditions. Participants' decision times for each task were measured in milliseconds from the moment the choice page was shown until a decision was made between option A, B, or delegation to AI.

Indecisiveness of participants was evaluated through the Indecisiveness Scale (IS), which originally consists of 15 items (e.g. "I try to put off making decisions"), measured on a five-point scale, ranging from *strongly disagree (1)* to *strongly agree (5)* (Frost & Shows, 1993). In the experiment, a shorter version, consisting of 11 items (Rassin, 2007) was used. Cronbach's alpha for this scale in the current study was .83, which indicates a high internal covariance.

Furthermore, we measured people’s gambling enjoyment using an adapted, shortened version of an existing scale (Lloyd et al., 2010). In the original scale, people are asked to reflect on 11 statements regarding the motivation for their gambling behaviour (e.g. “to relieve boredom”) on a 4-point Likert scale, including “never”, “occasionally”, “fairly often”, and “very often”. In the current study, only the six statements that measure enjoyment (as opposed to mood regulation and monetary incentives) were included. For the current study, a Cronbach’s alpha of .78 was found. Moreover, participants reported several demographics, which are described in the previous subsection (“Participants”).

Finally, several open questions regarding decision strategies were asked. For both contexts (hotel and coin gamble decisions), participants answered the following questions about four different decisions that they made during the experiment: “If you decide yourself, how would you make a choice between the two options?” and “Why would (or wouldn’t) you delegate this decision to AI?”. They also answered one final question: “Under which circumstances, in general, would you delegate a decision to AI and when would you make the decision yourself?”.

4.6 Procedure

The study was conducted between May 16th and June 4th, 2023. Participants completed the experiment online, using their own computers. After submitting the informed consent form, they encountered either the hotel or the coin context first and the other after that. Within the hotel context, an explanation of the situation and expectations was given and participants completed the point-allocation task and the matching tasks for this scenario as well as the decision-making tasks. The coin context contained an explanation, matching tasks, decision-making tasks and an evaluation of the bonus fee earned by the performance in the decision-making tasks.

After this, the demographics form followed, including age, gender, occupation, income, and whether one is a student. Participants filled in the indecisiveness and gambling enjoyment questionnaire and were asked some open questions about their decision strategies. Finally, participants were thanked for their participation, debriefed and some personal information needed for payment was submitted.

4.7 Analysis

STATA (StataCorp, 2021) was used for all analyses after pre-processing the data. To evaluate our hypotheses, multilevel logistic regression was used. To be more specific, we performed two-level logistic regressions with a random intercept at the participant level. The dependent variable is whether a decision is delegated or not. For each regression, the decision condition was included as the main independent variable, coded as a binary predictor variable with equally neutral as the reference category. Even though we did not have a hypothesis about this difference, we also ran the model with dominant as the reference category, to investigate potential differences between this condition and the trade-off 1 condition. All hypotheses were tested twice; once for the hotel context and once for the coin context. Adding to this, we ran a logistic regression model for all data combined as well. Moreover, various control variables were added to the models, including demographics, the indecisiveness and gambling enjoyment scales, decision times, count of choice trial, and absolute difference in exchange rate calculated with values from the point allocation task. Items from the indecisiveness scale were combined into one variable as well as the items from the gambling enjoyment scale. Decision time was log-transformed to ensure normality. The gender category “other” was added to the “no sex” category, as there was only one participant in the former.

Before running the models, a manipulation check was conducted as well as the necessary assumption checks. By calculating the percentages of participants that would choose either option A or B in each decision situation, we evaluated whether the stimuli were representative of the different conditions. The three equal conditions (good, neutral, and bad) as well as both trade-off conditions should have a distribution around 50/50 to indicate that the two available options are more or less equal in their subjective utility. The distribution of the dominant condition should be as close as possible to 100/0. If the percentages chosen in a dominant condition are smaller than 70 and bigger than 30 per cent, the dominant option is not dominant enough. If the percentages chosen in any of the other conditions are higher than 80 and smaller than 20 per cent, the rates are not close enough to 50/50. If these assumptions were not met, the models were run with and without including these stimuli to investigate the difference in results. Finally, the assumptions belonging to a mixed logistic regression (linearity and no multicollinearity between the independent variables) were checked and met.

To investigate the difference between the two contexts, patterns and differences between delegation rates in different decision conditions are qualitatively compared. Most importantly, we evaluated if there was a difference in which hypotheses were rejected and which were not. Furthermore, a two-sample test of proportions was conducted to see if there is a difference in delegation rates between the two decision contexts.

Finally, a thematic analysis as suggested by Braun and Clarke (2006) was conducted to analyse choice strategies and motivation to delegate decisions to AI. For both the hotel and the coin context, themes were identified for two open questions and the frequencies reported. The same was done for one final open question about delegation to AI in general.

5. Results

The results consist of multiple parts. First, the manipulation and assumption checks will be presented for both the hotel and the coin context. Then, the outcome of our logistic regression models, testing our hypotheses, will be explained for both contexts separately and the delegation rates of the two will be compared. Finally, the findings of the thematic analysis will be explained for both hotel and coin contexts separately as well as for delegation to AI in general.

5.1 Manipulation check

5.1.1 Hotel context

5.1.1.1 Point allocation task: exchange rate

For each of the pairs, different exchange rates were assumed based on pre-tests: for the price versus time pair, this was 3.5, for the cleanness versus bed comfort pair as well as the breakfast versus climate control pair it was 1. From the data, the following exchange rates, based on the point allocation task, followed: price versus time (M = 3.10, SD = 4.15, Mdn = 2.00), cleanness versus bed comfort (M = 2.26, SD = 3.09, Mdn = 1.35), breakfast versus climate control (M = 1.06, SD = .94, Mdn = 1.00).

The exchange rates in the cleanness versus bed comfort pair deviate most from what we expected, so extra attention was taken to this pair in further analysis. The other two pairs are close to what we expected. Moreover, as standard deviations are rather large, there is quite some variation between participants, so it is not certain that our stimuli are perfectly designed for everyone. For some, there might be a preference for one of the options, while there should not be, based on the design of the study. However, as there are no extreme deviations, we continued with the analysis as planned, but this notion is something to take into account.

5.1.1.2 Choice distributions

For the validation of our stimuli, the choice distributions for each of the stimuli in the three pairs were analysed (see Table 1-Table 3, p. 34). A stimulus that displays a perfect case of equal subjective utility would show a distribution of close to 50 per cent for option A and close to 50 per cent for option B. A perfect stimulus for a dominant case should display a percentage close to 100 for the dominant option and close to 0 for the other option according to utility theories.

When looking at Table 1, acceptable distributions are found for the price versus time pair. The same can be said about the breakfast versus climate control pair (Table 3). However, some of the distributions in the cleanness versus bed comfort pair (Table 2) have distributions that might be too extreme to be labelled as equal in subjective utility. For this reason, all analyses were performed with and without this pair[2] to find out how problematic the two imperfect stimuli are.

5.1.2 Coin context

5.1.2.1 Choice distributions

The choice distributions for all coin stimuli were also checked to confirm that the stimuli represent the correct circumstances (either a dominant option in the dominant condition or a more or less equal attractiveness for all other conditions). Almost all stimuli in Table 4 indeed show distributions close to what we expected, with some exceptions (trials 9, 11, 12, 13, and 16).

---

[2] Including or excluding the cleanness versus bed comfort pair (or even only the most "problematic" stimulus, trade-off 2) did not have a difference in terms of effect size. Because one third of the data was excluded by leaving this pair out, more p-values became larger than .05 than when all data was included, but the size and direction of all effects remained the same. For this reason, all data was left in all analyses from this point on.

As most stimuli are not too problematic, we continued with the analysis as planned, with a bit of extra caution regarding these stimuli.

5.2 Choice condition and delegation

5.2.1 Hotel context

Before investigating the effect of choice condition and several control variables on delegation to AI, assumptions were checked. Because of moderate to large correlations between age, occupation, and income, we only kept age as a predictor variable. Moreover, a quadratic relationship may be present between the log of decision time and delegation, but this cannot be said with certainty (see Appendix D, Figure 9, p. 81). Still, a new transformed variable was added to the model ($log(DT)^2$) to investigate this possible effect. When running the null-model of our logistic regression, we found *rho* = .471, which means that 47% of the variance in delegation rates is caused at the participant level.

With regards to the control variables that were added to the logistic regression model, no significant effects were found for indecisiveness, age, gender, whether they are a student, and the absolute difference between the expected exchange rate and personal exchange rate based on the point-allocation task. We did find a negative effect of the logarithmic function of decision time on delegation (*OR* = .02, $p < .001$, *95% CI* = [.00, .15]), a positive effect of the squared log of decision time on delegation (*OR* = 1.22, $p = .001$, *95% CI* = [1.08, 1.37]), and a slightly negative effect of the number of choice trials already encountered (*OR* = .96, $p = .006$, *95% CI* = [.93, .99]). The full logistic regression model can be found in Table 24, Appendix C (p. 78).

An overview of the choice distributions and delegation rates for all hotel stimuli is presented in Table 1, Table 2 and Table 3. We can see the lowest delegation rates for the conditions where there is one dominant option and the conditions where a large trade-off has to be made between the attributes. The highest delegation rates are found in situations where two options are equally bad.

Our logistic regression model (see Appendix C, p. 78) has given us reason to believe that delegation is indeed lower for the dominant condition compared to the equally neutral condition (*OR* = .45, p = .002*, 95% CI* = [.28, .75]). Delegation rates were also lower in the condition where a trade-off needed to be made between one balanced option and one option with extreme values (trade-off 1), compared to equally neutral (*OR* = .62, *p* = .059, *95% CI* = [.381, 1.018])[3]. When comparing the dominant to the trade-off 1 condition, we did not find evidence suggesting that there is a difference in delegation rate between these two conditions (*p* = .231). In addition, no support was found for the hypothesis that people delegate less often to AI in situations where a large trade-off has to be made in which one option is the "mirror image" of the other option, compared to situations when two equal options only involve a small trade-off (*p* = .294). With regards to the equally good, neutral, and bad conditions, no significant difference was found in the data between equally neutral and equally good (*p* = .876), but participants did seem to delegate more often in the equally bad condition, compared to equally neutral (*OR* = 1.71, *p* = .024, *95% CI* = [1.07, 2.73]). This means that we can confirm H1 and H2, reject H3 and reject H4 partially (equally good and neutral are the same, but bad is higher).

---

[3] Even though the p-value is larger than .05 here, we found the effect worth mentioning as we did not have clear a-priori predictions regarding what variables to and not to include in our model. When removing all control variables with p>.05 from the model, the p-value for comparting trade-off 1 with equally neutral is .048.

Table 1

Choice and Delegation Distributions for Each of the Stimuli in Pair 1 of the Hotel Context

Hotel	Price	Time	Percentage chosen	Percentage delegated
		Dominant		
Hotel A	€ 310,00	26 min.	3,08	18,75
Hotel B	€ 140,00	30 min.	96,92	
		Equally neutral		
Hotel A	€ 223,00	27 min.	43,64	31,25
Hotel B	€ 216,00	34 min.	56,36	
		Equally good		
Hotel A	€ 61,00	11 min.	43,40	33,75
Hotel B	€ 65,00	7 min.	56,60	
		Equally bad		
Hotel A	€ 388,00	54 min.	61,82	31,25
Hotel B	€ 380,00	62 min.	38,18	
		Trade-off 1		
Hotel A	€ 230,00	25 min.	52,73	31,25
Hotel B	€ 80,00	78 min.	47,27	
		Trade-off 2		
Hotel A	€ 360,00	8 min.	31,11	43,75
Hotel B	€ 55,00	73 min.	68,89	

Note: This table presents the stimuli in the hotel context where the attributes of interest are price (per night) and time (needed to travel to the city centre). It displays the distribution of participants choosing either option A or B, as well as the percentage of people who delegated this specific choice to AI.

Table 2

Choice and Delegation Distributions for Each of the Stimuli in Pair 2 of the Hotel Context

Hotel	Cleanness	Bed comfort	Percentage chosen	Percentage delegated
		Dominant		
Hotel A	5,7	6,6	3,39	26,25
Hotel B	6,2	8,1	96,61	
		Equally neutral		
Hotel A	6,9	6,3	69,81	33,75
Hotel B	6,5	6,7	30,19	
		Equally good		
Hotel A	8,6	9,1	32,73	31,25
Hotel B	8,9	8,8	67,27	
		Equally bad		
Hotel A	4,4	4,7	22,50	50,00
Hotel B	4,6	4,5	77,50	
		Trade-off 1		
Hotel A	6,4	6,6	39,06	20,00
Hotel B	8,9	4,5	60,49	
		Trade-off 2		
Hotel A	9,5	4,1	89,23	18,75
Hotel B	4,4	9,2	10,77	

Note: This table presents the stimuli in the hotel context where the attributes of interest are cleanness rating and bed comfort rating. It displays the distribution of participants choosing either option A or B, as well as the percentage of people who delegated this specific choice to AI.

Table 3

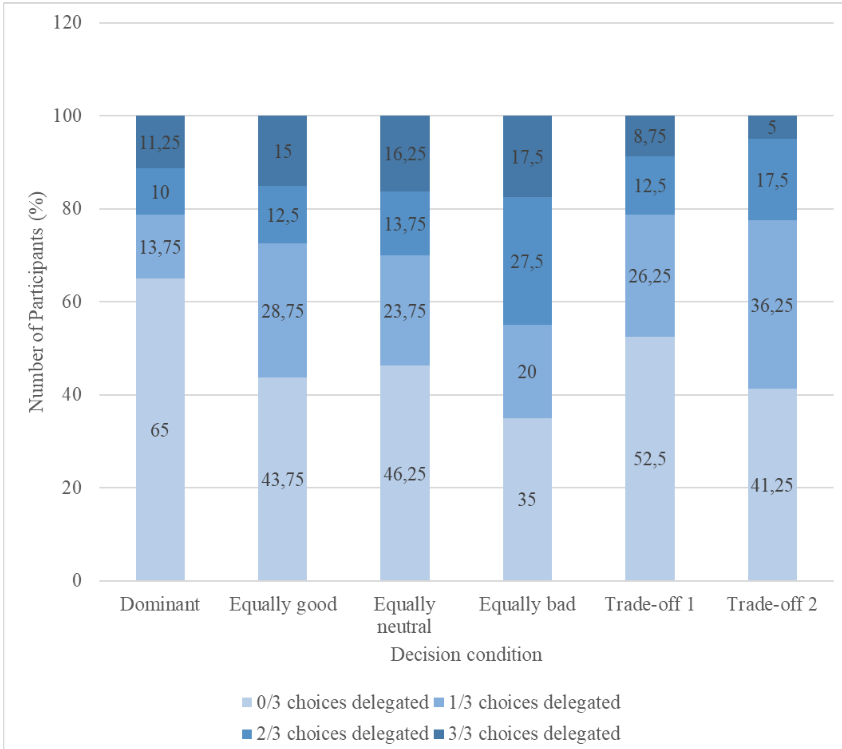
Choice and Delegation Distributions for Each of the Stimuli in Pair 3 of the Hotel Context

Hotel	Breakfast	Climate control	Percentage chosen	Percentage delegated
		Dominant		
Hotel A	8,6	5,9	85,48	22,50
Hotel B	5,4	6,2	14,52	
		Equally neutral		
Hotel A	6,6	6,4	53,85	35,00
Hotel B	6,2	6,9	46,15	
		Equally good		
Hotel A	9,0	8,7	62,26	33,75
Hotel B	8,6	9,2	37,74	
		Equally bad		
Hotel A	4,6	4,3	46,51	46,25
Hotel B	4,2	4,8	53,49	
		Trade-off 1		
Hotel A	6,7	6,3	59,32	26,25
Hotel B	5,0	8,7	40,68	
		Trade-off 2		
Hotel A	8,3	4,6	54,10	23,75
Hotel B	4,4	9,1	45,90	

Note: This table presents the stimuli in the hotel context where the attributes of interest are the rating for the breakfast offered and the rating for the climate control in the room. It displays the distribution of participants choosing either option A or B, as well as the percentage of people who delegated this specific choice to AI.

Figure 6

Number of Decisions Delegated per Participant per Condition in the Hotel Context



Note: This chart displays how many choices are delegated for each decision condition in the hotel context per participant. The percentages in the stacked bars present how many participants delegated none, one, two, or all three decisions in this condition.

As visually presented in Figure 6, there are many individual differences, even though in general we did find patterns and differences between the conditions. The total delegation rates for the dominant condition, for instance, are 18.75, 26.25, and 22.50 for the three pairs respectively (see Table 1-Table 3). However, Figure 6 shows that there are also participants who delegated all three of the decisions in this condition (11.25 per cent of all participants, which are 9 individuals), but at the same time there are participants who delegated none of the three decisions.

5.2.2 Coin context

In the second context, the coin context, the assumptions of a multilevel logistic regression model were all met. There are some small correlations between some of the variables, but nothing too extreme to take measures. A large part of the effects can be explained at the participant level (*rho* = .54).

Most control variables do not seem to have a significant effect on delegation (see Table 26, Appendix C, p. 79), except for age (*OR* = 1.05, *p* = .049, *95% CI* = [1.00, 1.10]) and the log decision time (*OR* = .66, *p* = .001, *95% CI* = [.51, .85]). However, when inspecting the effect of gamble enjoyment on delegation a bit more closely, a (non-linear) effect does seem to be present in the data (see Figure 12, Appendix E, p. 82). When adding the gamble enjoyment with a squared transformation to the model, however, still no significant effects were found in the data.

When inspecting Table 4, delegation rates appear to be lowest for the dominant condition, followed by trade-off 1. For the equally neutral, equally good, and trade-off 2 conditions, delegation levels seem to be comparable. Participants delegated most in de equally bad condition. Indeed, compared to the equally neutral condition, our model showed that delegation rates were lower in the dominant condition (*OR* = .21, *p* < .001, *95% CI* = [.12, .35]) and the trade-off 1 condition (*OR* = .42, *p* < .001, *95% CI* = [.26, .68]). When comparing trade-off 1 with dominant, we find higher delegation rates in the former than in the latter (*OR* = 2.02, *p* = .010, *95% CI* = [1.18, 3.43]). In the equally bad condition, delegation rates were significantly higher than in the equally neutral condition (*OR* = 2.35, *p* < .001, *95% CI* = [1.47, 3.75]). We did not find evidence for a difference in delegation when comparing the equally neutral condition with the equally good condition (*p* = .308) and with the trade-off 2 condition (*p* = .962). This means that we can, again, confirm H1 and H2, reject H3 and confirm H4 partially.

Table 4

Choice and Delegation Distributions for Each of the Stimuli of the Coin Context

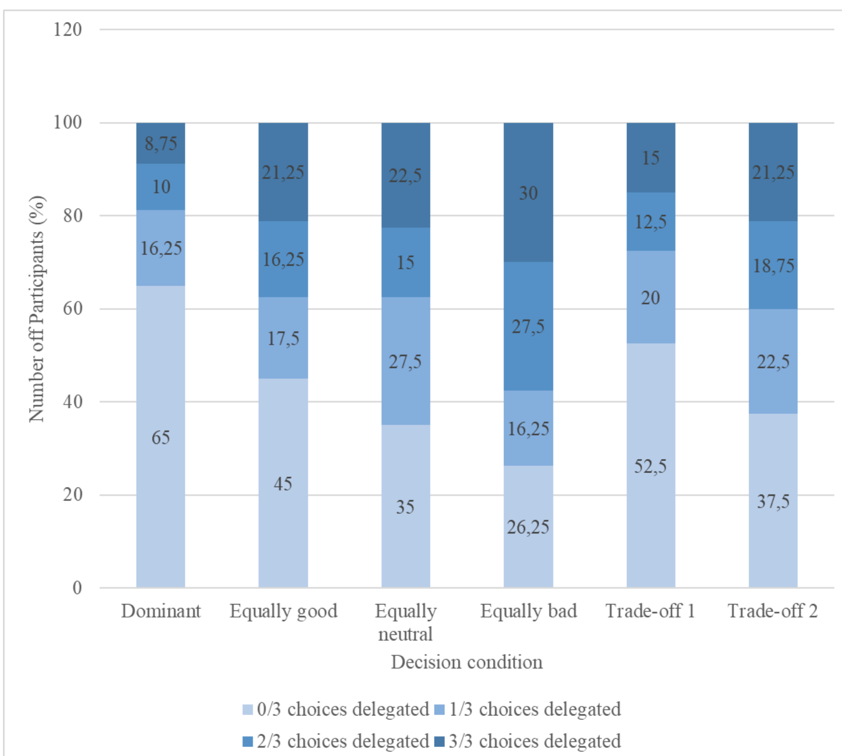
Trial	Option	Red coins	Blue coins	Percentage chosen	Percentage delegated
		Dominant			
1	Option A	-2,10	1,10	16,39	23,75
	Option B	-2,40	3,70	83,61	
2	Option A	-2,70	4,30	93,85	18,75
	Option B	-5,50	3,90	6,15	
3	Option A	-4,10	6,30	82,81	20,00
	Option B	2,60	-4,30	17,19	
		Equally neutral			
4	Option A	-0,40	0,50	42,55	41,25
	Option B	-0,30	0,40	57,45	
5	Option A	0,70	-0,90	43,18	45,00
	Option B	0,60	-0,80	56,82	
6	Option A	1,20	-0,90	65,31	38,75
	Option B	1,40	-1,10	34,69	
		Equally good			
7	Option A	4,70	3,40	46,00	37,50
	Option B	4,50	3,60	54,00	
8	Option A	5,70	7,30	40,38	35,00
	Option B	5,90	7,10	59,62	
9	Option A	8,40	9,10	29,79	41,25
	Option B	8,10	9,50	70,21	
		Equally bad			
10	Option A	-4,40	-3,40	48,57	56,25
	Option B	-4,30	-3,50	51,43	
11	Option A	-6,90	-6,10	21,05	52,50
	Option B	-6,60	-6,40	78,95	
12	Option A	-9,30	-8,70	18,42	52,50
	Option B	-9,10	-8,90	81,58	
		Trade-off 1			
13	Option A	4,30	-4,90	28,30	33,75
	Option B	-0,40	0,30	71,70	
14	Option A	7,60	-7,40	35,00	25,00
	Option B	0,60	-0,50	65,00	
15	Option A	-9,30	9,60	36,36	31,25
	Option B	0,50	-0,40	63,64	

Trial	Option	Red coins	Blue coins	Percentage chosen	Percentage delegated
			Trade-off 2		
16	Option A	5,10	-5,40	84,00	37,50
	Option B	-5,70	5,30	16,00	
17	Option A	8,50	-7,90	41,67	40,00
	Option B	-7,60	8,20	58,33	
18	Option A	-8,90	9,80	30,23	46,25
	Option B	9,60	-8,70	69,77	

Note: This table demonstrates the stimuli in the coin context. It displays the distributions of participants choosing either option A or B, as well as the percentage of people who delegated this specific choice to AI.

Figure 7

Number of Decisions Delegated per Participant per Condition in the Coin Context



Note: This chart displays how many choices are delegated for each decision condition in the coin context per participant. The percentages in the stacked bars present how many participants delegated none, one, two, or all three decisions in this condition.

Figure 7 shows that individual differences in delegation are large in the coin context, too. The overall patterns are visible when looking at the total population and indications of these patterns are visible in this figure as well, but it is important to notice the variance in the number of choices that each participant delegates.

5.2.3 Overall

When looking at Table 1 (p. 34), Table 2 (p. 35), Table 3 (p. 36), and Table 4 (p. 39), overall delegation rates seem to be lower for the hotel context compared to the coin context. When combining the data from all trials, this is still the case (see Table 5). A two-sample test of proportions indeed showed a significant difference in delegation rates between the two contexts ($p < .001$).

For both contexts, the same hypotheses were confirmed and rejected. Moreover, this is still the case when adding all data from both contexts in a logistic regression together (see Table 28, Appendix C, p. 78). We have reason to believe that delegation is lower for situations where there is one dominant alternative or when there is a trade-off between one option that has more balanced values and one option with more extreme values, compared to a situation where there is only a small difference between all relatively good or neutral values (H1 and H2). We did not find evidence for the hypothesis that delegation is higher in cases where a large trade-off has to be made between two options that are the "mirror image" of each other compared to situations with a very small trade-off (H3). When a choice has to be made between two equally bad options, delegation is higher than choices between two equally good or neutral options, which partially rejects our final hypothesis (H4).

Table 5

Total Delegation Rates for Both the Hotel and the Coin Context

Choice condition	Delegation hotel context (%)	Delegation coin context (%)
Dominant	22.50	20.83
Equally neutral	33.33	41.67
Equally good	32.92	37.92
Equally bad	42.5	53.75
Trade-off 1	25.83	30.00
Trade-off 2	28.75	41.24

Note: This table demonstrates the delegation rates per context of all 18 trials combined.

5.3 Thematic analysis

For the final part, open questions were asked to understand people's decision strategies and motivations to delegate to AI a little better. We randomly selected four trials from the choice tasks: one from equally good, one from equally bad, and one from each of the two kinds of trade-offs. The values of the choice were shown and we specified whether the participant had delegated this specific decision or not. Regardless of their previous decision in this situation, the same two questions about the choice were asked to all participants: "If you decide yourself, how would you make a choice between the two options? What are the things that you take into consideration?" and "Why would (or wouldn't) you delegate this decision to AI?" From the answers, codes were extracted and themes were created.

5.3.1 Hotel context

The following themes were extracted from the answers to the question about which things are taken into consideration when making a decision without the help of AI:

- Prominence: one of the two attributes is strongly favoured over the other;

- Min/max constraints: the person has an upper or lower limit to the value of one of the attributes, regardless of the other;
- Mathematical reasoning: the values of the two options are compared and the decision is based on this comparison;
- No preference: people are indifferent or uncertain about which option to choose;
- Other properties: people state that they would look at other attributes, details or circumstances than the ones provided in the table to make this decision.

See Table 6 for an example quote for each of the identified themes.

Table 6

Themes in Open Question 1 (Hotel Context)

Theme	Quote
Prominence	"I care more about bed comfort than cleanness, so I'll take Hotel A." (participant 3)
Min/max constraints	"73 minutes is way too long of a travel time for me so in this case I wouldn't mind spending more money" (participant 42)
Mathematical reasoning	"I look at what has the "highest" scores respectively. Overall this is the case for hotel A, where the lowest score is 4.3, compared to the 4.2 of hotel B." (participant 18)
No preference	"Probably completely random, because I find both hotels not attractive at all. Cleanness and bed comfort are important to me and both score low on those characteristics." (participant 6)
Other properties	"What the impression of both hotels are" (participant 66)
	"I would look at all the other characteristics and base my decision on that." (participant 6)

Note: Example quotations from each theme defined from the answers to the question "If you decide yourself, how would you make a choice between the two options? What are the things that you take into consideration?"

Noted first is that prominence is named most frequently as a motivation to choose one of the two options. This is the case for all four conditions. Min/max constraints are present more often in the trade-off conditions than in the equal conditions, while at the same time, people indicate to be having no preference more often in the equal conditions than in the trade-off conditions.

Table 7

Frequencies of Occurrences of Themes in Open Question 1 (Hotel Context)

Theme	Equally good (*f*)	Equally bad (*f*)	Trade-off 1(*f*)	Trade-off 2 (*f*)
Prominence	40	45	48	49
Min/max constraints	3	2	14	12
Mathematical reasoning	13	9	15	7
No preference	24	19	2	6
Other properties	2	5	2	4

Note: The answers to the question "If you decide yourself, how would you make a choice between the two options? What are the things that you take into consideration?" were analysed and extracted themes were counted and presented in this table.

The themes (as seen in Table 8) identified as reasons to delegate to AI are:

- No preference: People are indifferent or uncertain about which option to choose;
- Curiosity: People are curious about the decision of the AI;
- Trust in AI: People trust in terms of calculation abilities or in terms of representing their personal preferences.

The themes identified as reasons not to delegate the decision to AI (Table 9) are:

- No trust in AI: AI is not trusted in making the preferred decision or the person does not believe the AI is trained well enough by a limited amount of data;
- Maintain control: People prefer to be in control and thus responsible for the outcome;

- Clear preference: One option is (strongly) favoured over the other;
- High risk: There is a lot at stake for the participant or the risk of a bad outcome is too high.

Table 8

Themes in Open Question 2, in Favour of Delegation (Hotel Context)

Theme	Quote
No preference	"I would delegate this decision to AI because I don't find either breakfast or climate control particularly important" (participant 49)
	"Because I was definitely not sure about this choice because of the big differences." (participant 23)
Curiosity	"I delegated since I was curious what the ai would pick" (participant 68)
Trust in AI	"Because it can calculate my trivial logic more easily than me. These are really simple calculation problems, I don't want to do them myself." (participant 8)

Note: Example quotations from each theme defined from the answers to the question "Why would (or wouldn't) you delegate this decision to AI ?"

Table 9

Themes in Open Question 2, Against Delegation (Hotel Context)

Theme	Quote
No trust in AI	"because AI doesn't take my maximum budget into consideration" (participant 15)
Maintain control	"It's not a huge decision and I just like to make the decision on my own." (participant 30)
	"I wouldn't delegate this decision, mainly because I don't feel the need to delegate. I am perfectly capable of making this choice on my own without hesitation" (participant 65)
Clear preference	"Because I already knew I rather have a better bed." (participant 19)
High risk	"The difference between the two is too big" (participant 43)

Note: Example quotations from each theme defined from the answers to the question "Why would (or wouldn't) you delegate this decision to AI ?"

Again, it is visible that participants had no preference between the options in the equal conditions more often than in the trade-off conditions (Table 10). However, people indicated to have a clear preference between options approximately equally often in the equally good and both trade-off conditions, only somewhat less often in the equally bad condition (Table 11). Also, there is a lack of trust in AI more often for the trade-off conditions (especially trade-off 2) than for the equal conditions. Interesting to note, but not visible in the tables, is that some people mentioned that they delegated some decisions in the experiment out of curiosity, which they would not delegate in real life.

Table 10

Frequencies of Occurrences of Positive Themes in Open Question 2 (Hotel Context)

Theme	Equally good (*f*)	Equally bad (*f*)	Trade-off 1(*f*)	Trade-off 2 (*f*)
No preference	33	37	17	6
Curiosity	3	2	3	3
Trust in AI	1	3	3	5

Note: The answers to the question “Why would (or wouldn’t) you delegate this decision to AI ?” were analysed and extracted themes in favour of delegation were counted and presented in this table.

Table 11

Frequencies of Occurrences of Negative Themes in Open Question 2 (Hotel Context)

Theme	Equally good (*f*)	Equally bad (*f*)	Trade-off 1(*f*)	Trade-off 2 (*f*)
No trust in AI	8	8	11	18
Maintain control	4	3	5	8
Clear preference	20	12	26	22
High risks	0	2	4	7

Note: The answers to the question “Why would (or wouldn’t) you delegate this decision to AI ?” were analysed and extracted themes against delegation were counted and presented in this table.

5.3.2 Coin context

Again, four trials from the choice tasks were presented to the participants and the following question was asked about each of them: “If you decide yourself, how would you make a choice between the two options? What are the things that you take into consideration?” Themes were created from the answers and summarized in Table 12:

- No preference: the two options are deemed equal or a random choice is made between the two options;
- Maximize profits: the option with the highest value is chosen, more risk is taken in trade-off situations, or people mention to be “testing their luck”;
- Minimize loss: the option with the least negative value or the risk-averse option in trade-offs is chosen, or one prefers the option where the difference between the blue and red coins is smaller.
- Calculated preference: a participant did not specify a preference for the highest possible win or lowest possible loss, but did take into account the difference between values;
- Intuitive preference: A preference based on feeling or intuition, or without any specified underlying reason.

The most present reasons for making decisions are having no clear preference for any of the two options (while this reason is smaller in the trade-off 1 condition) and wanting to minimize the potential loss. Other differences and how these differences could explain the outcomes of our quantitative data analysis will be discussed in the discussion section (p. 54).

Table 12

Themes in Open Question 1 (Coin Context)

Theme	Quote
No preference	"No idea. Random guess." (participant 24)
Maximize profits	"the chance to win big can make it worth it" (participant 37)
Minimize loss	"I always want to lose less regardless of gain, so I would pick B." (participant 9)
Calculated preference	"The absolute difference I would win/lose on average" (participant 74)
Intuitive preference	"I would choose option B because it looks more attractive to me than option A." (participant 36)

Note: Example quotations from each theme defined from the answers to the question "If you decide yourself, how would you make a choice between the two options? What are the things that you take into consideration?"

Table 13

Frequencies of Occurrences of Themes in Open Question 1 (Coin Context)

Theme	Equally good (*f*)	Equally bad (*f*)	Trade-off 1(*f*)	Trade-off 2 (*f*)
No preference	28	35	10	31
Maximize profits	16	0	12	8
Minimize loss	5	26	39	24
Calculated preference	12	5	4	3
Intuitive preference	3	4	2	3

Note: The answers to the question "If you decide yourself, how would you make a choice between the two options? What are the things that you take into consideration?" were analysed and extracted themes were counted and presented in this table.

The second question we asked about the selected four trials was "Why would (or wouldn't) you delegate this decision to AI?" The identified themes in favour of delegation can be found in Table 14:

- No preference: indifference or uncertainty about which option to choose;
- Curiosity: curious about what AI would choose;
- Trust in AI: trust in terms of calculation capabilities or trust in the algorithm that trained it based on personal preferences;
- Avoid responsibility: not wanting to choose to prevent a bad outcome or not wanting to gamble at all.

The themes for no delegation can be found in Table 15:

- No trust in AI: being afraid the AI will choose wrong or thinking it does not know the personal preferences well enough;
- Maintain control: wanting to decide by themselves;
- Clear preference: an easy decision which does not need the help of AI or being so certain of a choice that it is better/safer to just choose instead of delegate.

Table 14

Themes in Open Question 2, in Favour of Delegation (Coin Context)

Theme	Quote
No preference	"I would delegate this decision to AI because it is a choice without many risks. If it chooses the wrong option, it's not terrible." (participant 36)
Curiosity	"I was curious to see whether the AI would take the more risk averse or risk taking option" (participant 74)
Trust in AI	"AI would probably make the same choice as me." (participant 22) "Trust AI to be smarter than me when it comes to numbers, and I had a hard time figuring out which one was better. So might as well have the AI figure it out." (participant 29)
Avoid responsibility	"I would let ai make this decision since id be too afraid to take the risk myself" (participant 16)

Note: Example quotations from each theme defined from the answers to the question "Why would (or wouldn't) you delegate this decision to AI?"

Table 15

Themes in Open Question 2, Against Delegation (Coin Context)

Theme	Quote
No trust in AI	"I wouldn't delegate it because I would not trust the AI to make a good decision" (participant 28)
	"I don't trust the AI trained on my imperfect math skills to be better at math than me" (participant 48)
Maintain control	"It's the game for testing my luck not for AI. That's I would like to make the choice on my own." (participant 30)
Clear preference	"I did not have to think long over this so I did not need the AI" (participant 33)

Note: Example quotations from each theme defined from the answers to the question "Why would (or wouldn't) you delegate this decision to AI?"

The reason that was named most often as a reason to delegate is that there is no preference between the options. This reason is most named in all conditions, but the least in the trade-off 1 condition. As reasons not to delegate, people mention various things, with the highest overall frequencies in the trade-off 1 condition for all possible reasons.

Table 16

Frequencies of Occurrences of Positive Themes in Open Question 2 (Coin Context)

Theme	Equally good (*f*)	Equally bad (*f*)	Trade-off 1(*f*)	Trade-off 2 (*f*)
No preference	42	40	13	37
Curiosity	4	2	7	5
Trust in AI	6	3	5	2
Avoid responsibility	2	5	1	3

Note: The answers to the question "Why would (or wouldn't) you delegate this decision to AI?" were analysed and extracted themes in favour of delegation were counted and presented in this table.

Table 17

Frequencies of Occurrences of Negative Themes in Open Question 2 (Coin Context)

Theme	Equally good (*f*)	Equally bad (*f*)	Trade-off 1(*f*)	Trade-off 2 (*f*)
No trust in AI	8	10	24	11
Maintain control	7	9	11	8
Clear preference	5	4	10	5

Note: The answers to the question "Why would (or wouldn't) you delegate this decision to AI ?" were analysed and extracted themes against delegation were counted and presented in this table.

5.3.3 Overall

We asked participants one final question, regarding their motivation to delegate or not: "Under which circumstances, in general, would you delegate a decision to AI and when would you make the decision yourself?" Not every participant answered both parts of the question. Some, for instance, only answered when they would not delegate, which is why the numbers in the columns of Table 20 are not equal.

As reasons to delegate, these themes were mentioned (Table 18):

- Difficult: The decision is too difficult to make or there is no preference between the options;
- Low stakes: The outcome does not matter too much;
- Low risk: There is not a lot to lose in this decision;
- Small differences: Both options are similar, the differences are not that big;
- Objective topic: In this choice, mathematics or probability is involved;
- AI validated: The participant is certain of the algorithm used by the AI.

The reasons against delegation are the following (Table 19):

- Easy: There is a clear dominant or preferred option;

- High stakes: The outcome is important;
- High risk: There is too much to lose;
- Large differences: The options are very different from each other;
- Subjective topic: The decision is about a personal topic or something where people have individual differences;
- No trust in AI: The algorithm is not validated or trusted. People do not believe the AI would make the right decision.

Table 18

Themes in Open Question 3, in Favour of Delegation

Theme	Quote
Difficult	"I would delegate a decision to AI when I see obvious tradeoffs that I find tough." (participant 1)
Low stakes	"If it's an unimportant decision, AI can make it for me." (participant 3)
Low risk	"If […] the risk isn't to high I'd consider AI" (participant 51)
Small differences	"In general, I would choose AI if there is little difference between the options." (participant 17)
Objective topic	"I would delegate options to AI when it comes to deciding the value of one choice over the other when it comes to making money or deciding prices." (participant 37)
	"In mathematical/statistical/probability problems, I would easily let AI take the lead in deciding for me." (participant 6)
AI validated	"if I have checked the underlying algorithm and the accuracy/recall/F1 score of the AI i would trust it more" (participant 10)

Note: Example quotations from each theme defined from the answers to the question "Under which circumstances, in general, would you delegate a decision to AI and when would you make the decision yourself?"

Table 19

Themes in Open Question 3, Against Delegation

Theme	Quote
Easy	"I would make the decision myself when it seems easy, and I know exactly what I prefer." (participant 1)
High stakes	"If something larger is at stake, I'd like to take matters into my own hands." (participant 3)
High risk	"When the risk is higher, I would rather make the choice myself, especially for choosing hotels." (participant 4)
Large differences	"If there are significant differences, then I prefer to decide for myself." (participant 17)
Subjective topic	"When talking about more abstract categories such as cleanness or breakfast I can decide for myself what category I rank above others and decide the relative value of different scores on these scales myself." (participant 37)
No trust in AI	"Here, I knew the AI was only trained on a few questions, so I did not trust it to follow my way of thinking" (participant 28)

Note: Example quotations from each theme defined from the answers to the question "Under which circumstances, in general, would you delegate a decision to AI and when would you make the decision yourself?"

The results presented in Table 20 show that the decision displaying small differences is the most important reason to delegate a decision to AI. Also interesting to note is that indeed opposite reasons are given for delegation versus no delegation. Every theme has a counterpart in the other column. Worth mentioning is that some (three) people said that they would not delegate decisions to AI, but that they would take advice from AI, as long as they have the final say.

Table 20

Frequencies of Occurrences of Themes in Open Question 3

Reason to delegate	Frequency	Reason not to delegate	Frequency
Difficult	18	Easy	11
Low stakes	22	High stakes	15
Low risk	5	High risk	3
Small differences	32	Large differences	9
Objective topic	5	Subjective topic	8
AI validated	5	No trust in AI	8

Note: The answers to the question "Under which circumstances, in general, would you delegate a decision to AI and when would you make the decision yourself?" were analysed and extracted themes were counted and presented in this table.

6. Discussion

This study was conducted to get more insights into the decision to delegate a choice to AI in different decision situations. More specifically, we wanted to investigate how choice characteristics influence people's willingness to delegate to AI in different decision contexts. As this study is partly a replication of van den Bergen (2023), we expected to find similar results as she did. Our main hypothesis was that difficult decisions are delegated more often than easy ones. For our decision conditions, this meant that we expected the following pattern: choices in which one dominant alternative is present are delegated least often, as well as choices with a trade-off between one balanced option and one option with extreme values; delegation rates are moderate for choices between two options that are equal in subjective utility; we did not predict a difference between situations where options are equally good, equally bad, or neutral; choices in which there is a trade-off in which one option is a cross-over of the other options are delegated most often.

6.1 Interpretation of results

As explained earlier, we made several changes compared to the study by van den Bergen (2023) to address some unanswered questions. The main goal was to make the two contexts more comparable in terms of how the values of the stimuli are presented to check if the differences between the contexts are caused by the difference in setting or by the difference in stimuli presentation. The initial conclusion is that we succeeded in our goal, as we did not find that certain hypotheses were confirmed in one context but rejected in the other, while van den Bergen (2023) did find this difference between the contexts. Overall delegation rates were still higher in the coin context than in the hotel context, which can be explained by the fact that people more often have a personal preference in more subjective, real-life situations (such as choosing a hotel) compared to situations where chance and numbers play a larger role (such as a coin gamble game). This was also one of the findings of our thematic analysis.

6.1.1 Equal versus dominant condition

Our findings suggest that our first hypothesis (H1: People delegate less often to AI in situations where one choice option dominates the other, compared to situations where the two options are equal (or close to equal) in terms of their subjective utilities) is likely to be true in both the hotel and the coin context. This is in line with our general assumption that delegation is lower for easier decisions.

6.1.2 Equal versus trade-off (balanced versus extreme) condition

Our second hypothesis (H2: People delegate less often to AI in situations where a large trade-off has to be made in which there is one option with extreme values for both attributes and one balanced option, compared to situations when two equal options only involve a small trade-off) was also supported by our results. People delegated less often in the trade-off 1 condition, compared to the equally good, equally neutral, and equally bad conditions. This is partly in line with previous findings by van den Bergen (2023). Where she only found support for this hypothesis in the coin context, we found results supporting it in both contexts. This can be explained by the fact that, in the current study, the stimuli values in the hotel context were changed to resemble the coin context more. By changing the range from 0-7 to 0-10 and making the prices and travel times more extreme, the scores became more intuitively interpretable and the difference between moderately sufficient values (in the balanced options) and very good or very bad values (in the extreme options) became more apparent.

When looking at our qualitative data, we find that for the hotel context, the most important aspect taken into consideration when choosing without the involvement of AI is the prominence of one of the attributes. However, as this is the case for all conditions, this does not aid in finding an explanation for the fact that delegation is higher for equal conditions compared to the trade-off with a balanced option. We did find that people have constraints in their minimal or maximal requirements for a hotel more often when encountering a trade-off than when options are equal and that they do not have a preference between options more often in the latter case. This could indicate that for the trade-off condition, people can justify their decision, which makes the decision easier and delegation less favourable. If we connect this to the inverted U-shape of the double-mediation model by Scholten and Sherman (2006), we can conclude that our trade-off 1 condition displays a trade-off that is so large that choices become easy again, rather than moderately large which would have resulted in more difficult choices because of the sacrifices that need to be made.

For the coin context, our qualitative data show more instances of “no preference” in the equal conditions compared to trade-off 1 again. We also found that, in the trade-off condition, more participants filled in that they want to minimize their loss. When comparing to equally good, this is a logical effect, since there are no losses involved in this condition. However, there are also more instances of people wanting to minimize their loss in trade-off 1 compared to the equally bad condition, which could explain why they want to stay in control in the trade-off condition and will delegate less as a result. When asking participants why they would or would not delegate their decision to AI, it became clear that people have a clear preference for one of the two options most often in the balanced trade-off condition and do not trust AI in making the right decision more often in this condition than in the other. The preference in combination with a potential lack of trust can explain the lower delegation rates in this context, too. As the majority of participants chose the more balanced option over the option with extreme high and low values, we are tempted to think that the main reason for this effect for most people is loss aversion; the risk to lose a lot weighs more than the chance to win a lot.

6.1.3 Equal versus crossover trade-off condition

We did not find indications in our data to support our third hypothesis (H3: People delegate more often to AI in situations where a large trade-off has to be made in which one option is the “mirror image” of the other option, compared to situations when two equal options only involve a small trade-off) in either of the contexts. We tried to relate this to the inverted U-shape of the double-mediation model again (Scholten & Sherman, 2006). Note that this model shows how difficulty is low for both very small and very large trade-offs, but higher for moderate trade-offs. However, it is important to establish what people regard as a small, moderate, or large trade-off exactly. One might think that our trade-off 2 condition displays a very large trade-off and our equal conditions a very small one, but it is very well possible that consumers do not see it that way. In a trade-off between one option with more balanced values and one option with more extreme values (as in our trade-off 1 condition), it is reasonable to believe that the perceived similarity is very low, which might result in a very high perceived trade-off size. However, when two options are each other's mirror image, the perceived similarity might be higher and the trade-off size smaller. While the absolute difference between the values may be larger, the perceived trade-off size may not necessarily be. This could mean that our equal conditions and our trade-off 2 condition have a comparable level of perceived similarity and are thus equally difficult, while our trade-off 1 condition is easier, resulting in lower delegation rates.

Contrary to our results, van den Bergen (2023) did find that delegation rates were higher in the cross-over trade-off condition than in the equal conditions, but only for the coin context. In her case, the options in the trade-off 2 condition were an exact mirror image of each other. If one could choose 5,60 red coins and 4,20 blue ones in option A, option B would be 4,20 red coins and 5,60 blue ones. As, in the present research, we changed the stimuli to be slightly different from each other, one option became more balanced than the other. By doing this, we created a situation that is more comparable to our other trade-off condition. Delegation rates in the trade-off 1 condition are lower than in the equal conditions, which can explain the fact that delegation rates in the trade-off 2 condition are not significantly higher than in the equal conditions (anymore). Furthermore, as Kim et al. (2013) already found, the perceived similarity between two options increases when slight differences within the attributes are present. As people regard the two options as more similar, they might also see it as a smaller trade-off than when the options are an exact cross-over of each other, which also bring the condition more to the middle of the inverted U-shape of the double-mediation model (Scholten & Sherman, 2006) and more towards being comparable to our equal conditions. Note that we also changed the rule in the coin scenario from one of the coins doubling in value to one of the coins becoming worthless. While it is not logical to assume that this change caused the above-described difference in results compared to what van den Bergen (2023) found, we should keep this in mind.

6.1.4 Equally good/neutral versus equally bad condition

The results of our logistic regressions give us reason to conclude that delegation rates are higher for situations where two options are both equally bad compared to situations where options are both really good or neither good nor bad. This means that we have to partially reject the fourth hypothesis (H4: When two options are equal (or close to equal) in terms of their subjective utilities, the overall attractiveness of the two options does not influence the likelihood of delegation to AI (no differences among equally good, bad, and neutral conditions)). We hypothesised that there would be no difference as a consequence of the findings by van den Bergen (2023). She also believed there would not be a difference between equally good and bad conditions because earlier research had shown equal delegation rates for losses and gains (Candrian & Scherer, 2022). However, when assuming a positive relation between decision difficulty and delegation, it makes sense to assume a higher delegation for decisions with only bad alternatives as conflicts on negative attributes are rated as more difficult than conflicts on positive attributes (Chatterjee & Heath, 1996). The difference with van den Bergen (2023) could be caused by the fact that we made the values of all stimuli somewhat more extreme. So the equally good conditions became even better, but more importantly, the equally bad conditions became even worse. Increasing the contrast between the conditions might have resulted in a bigger difference in difficulty and thus in delegation.

Interesting to note is that when examining the delegation rates of each of the attribute pairs in the hotel context, delegation does not seem to be higher in the equally bad condition compared to equally neutral and equally good in the price versus time pair. This could be because there is not an almost objective qualitative attractiveness to price and time, while there is for ratings on a scale of 0 to 10. Expensive can always get more expensive and travelling time can always get larger, while a score of 4.5 or lower is a bad rating, regardless of whom you ask.

The thematic analysis of the hotel context does not very clearly give a simple explanation for the difference between the equal conditions. People mention having no preference for both contexts approximately equally often. However, as a reason for not delegating a decision, a clear preference for one of the two options was named more often in the equally good than in the equally bad condition. The themes in the coin context also show this difference to some extent. More people say to have a calculated preference for either of the options in the equally good condition compared to the equally bad condition when explaining the things they take into consideration when making the choice themselves. However, when asking why people would or would not delegate the decision, there is again no difference in how often people do or do not have a preference between the options in equally good versus equally bad conditions.

6.1.5 Additional factors that influence delegation

In both contexts, a negative relation was found between the log transformation of decision time and delegation. In the hotel context, one might believe that this relation is even quadratic (see Figure 9, Appendix D, p. 81), but after closer inspection, it is not possible to assume this quadratic relation. The number of data points for which the curve goes up again is a lot sparser than in the decreasing part of the curve. Moreover, decision times for which the delegation rate goes up again are very high. We measured the decision time from the moment the stimulus is presented to the moment a decision is made. However, it is very well possible that these times are not the actual decision times, but a period in which participants also engaged in other activities besides evaluating the options and choosing between making a decision or delegating. A participant might, for instance, have the experiment opened on their computer, but went to get a cup of coffee in the meantime.

If we conclude that there is a negative relationship between decision time and delegation for both contexts, we can interpret this as follows: People scan the option first and decide to delegate if they do not have an immediate preference or other reason not to delegate. If they take more time, they evaluate the options in more detail and develop a preference for either of the two options. Once this preference or justification is present, they do not delegate but make the decision themselves. However, as we merely measure the time it takes from the moment the stimulus is presented to the moment a choice is made, these interpretations are only speculations and no real theories.

Only in the hotel context, the trial count (how many trials out of the 18 were already encountered before this one) seems to have a slightly negative effect on delegation. We do not have a very well-reasoned explanation for this, but it could be the case that participants found that the AI made the wrong decision on several occasions, which made them not want to delegate anymore. In the coin context, personal preference for either of the options might be lower, which causes the lack of an effect in this context.

When investigating the effect of gamble enjoyment on delegation in the coin context, our logistic regression models did not give reason to believe there is a relation between these two variables. However, when plotting the data (see Figure 12, Appendix E, p. 82), some patterns seem to be there. Gambling can work in two different ways. One might see delegation as a gamble in itself because the outcome is unknown. At the same time, when a person enjoys gambling, they might prefer to actually perform the gamble over letting someone else (or something else in the case of AI) do it for them. That could be an explanation for the fluctuating pattern seen in Figure 12. However, these thoughts are merely speculations rather than well-reasoned theories.

In the coin context, our data indicate a positive effect of age on delegation, although not a very large one. We did not find a possible cause for this effect, however. In addition, one might expect a positive relationship between the indecisiveness trait of a person and their willingness to delegate decisions, but we did not find such a relation in our data.

6.2 Limitations

Several limitations can be identified for this study. As with almost all research, our stimuli were not a perfect presentation of reality. In reality, decisions are based on many more than just two attributes and often decisions have to be made between more than just two options. This also holds for decisions where AI or ADM could be deployed. As having more options to choose from increases choice difficulty (Scheibehenne et al., 2010), this could also influence decision strategies and, in turn, motivations to delegate and delegation rates. Furthermore, in real life, not all the needed information regarding a decision and its circumstances is known to the decision-maker. This also makes our experimental setting different from the real world. In the qualitative analysis of the hotel context we also found that some people mentioned that they delegated some decisions in the experiment out of curiosity, which they would not delegate in real life. To really be able to draw practically relevant conclusions from our findings, more realistic, real-life situations need to be investigated first.

Another limitation concerns the hotel context. When looking at the choice distributions, not all stimuli seemed to display the assumptions that we made regarding the decision contexts in terms of comparable attractiveness. Especially in the cleanness versus bed comfort pair, many participants seemed to have a clear preference for one option over the other (more often than not, they had a preference for the cleaner hotel). Moreover, our thematic analysis showed high frequencies for prominence as motivation to choose between the options. This can be an indication that there is indeed no subjective equal utility for all stimuli for all participants. Combining these notions illustrates the importance of personally tailored stimuli. Even for the stimuli where the choice distributions were around 50 per cent for both options, one option might still have been more attractive than the other for certain individuals.

Furthermore, while the thematic analysis did provide us with useful insights into when and why people are willing to delegate and when they are not, the way of analysing the data does bring some restrictions with it. By merely exploring the frequencies of the different themes over the entire sample, the analyses average over participants of which some might delegate a lot and some do not. Classifying the participants in some way (e.g. their personal delegation rate) and then looking at patterns for delegation motivation could provide very valuable additional insights, taking into account individual differences in overall willingness to delegate.

6.3 Future research

Concluding from the described limitations, the first direction for future research is using personalized stimuli so that the exchange rate of individuals is taken into account in the decision tasks. This can ensure that all stimuli are equal in subjective utility for all participants. As van den Bergen (2023) already mentioned in her discussion, it is recommended to use additional or other methods to measure personal exchange rates between attributes. Examples of this are conjoint analysis (Louviere & Islam, 2008) or training a latent-feature diversification algorithm (Willemsen et al., 2016). Also, more personal differences might be taken into account and investigated more in-depth. For example, gamble enjoyment (but in more detail than has been done now), people's knowledge about and experience with AI or technology in general, or expertise and confidence about the topic.

A second recommendation for future research is to compare delegation rates to AI to delegation rates to a human agent. As described in our related work section, current research comparing delegation to AI versus humans is not consistent and no research into how choice characteristics influence this difference is conducted yet, to our knowledge. Especially interesting would be to investigate this difference in a new, socially relevant context (e.g. hiring a new employee). It could be insightful to see whether people prefer to delegate these decisions that have actual consequences for other humans to AI, a colleague, or not at all.

A third direction for future research is investigating the mental model people have about AI and how this influences delegation. One might investigate how the transparency or accuracy of the algorithm of the decision-making AI influences delegation. In the current study, participants were only told that the AI was being trained based on their preferences, which they indicated in a (small) number of tasks. We did not validate this manipulation, which means we cannot be certain that participants believed the "realness" of the AI. If, in a future study, we would measure how accurately people believe the AI can represent their preferences, this could be used in the analysis and provide a deeper understanding into how transparent a system should be.

Finally, someone might be interested in researching what people expect from or want in an artificially intelligent agent. The assumption is made in the current study that an AI trained on personal preferences is the type of AI that people would delegate a decision to, but this might not be the case. Especially for objective contexts (like the coin context in this research), people may want an AI that can make the optimal, rational, calculated choice. On the other hand, it could be possible that for emotional or social contexts, someone might prefer an AI that makes human-like or even almost emotional (sometimes somewhat irrational) decisions. When we know what users expect or want from AI, this can be used to design systems that will be enjoyed and used optimally.

6.4 Practical implications

As our study provides insights into under which circumstances people are more willing to delegate to AI, these findings could be translated to how we can deploy AI in decision-making more effectively. It seems like decision difficulty and delegation are related, as we expected, but there is more to it.

In general, we can assume that is safe to utilize AI in decision-making when stakes are not too high and differences are not too big, especially in situations where people do not have personal preferences. Most people are willing to delegate decisions in which the options are fairly equal and the outcome does not matter too much. While this seems like a less useful area to use AI (as the outcome does not matter anyway), it can still save a lot of manual work and time if we can be relatively certain that people are willing to hand over the responsibility for these decisions.

However, there are still many people who do not trust AI in making the right decision in a lot of cases. When decisions do not matter, they might be willing to delegate, but as soon as something bigger is at stake they want to keep control over the outcome of the decision. This, rather than investigating when AI is most useful, might be a more pressing problem to solve and do more research into. When more is known about reasons for trusting or not trusting AI, delegation rates might be increased and usage of these systems enhanced.

7. Conclusion

This study provides new insights into how choice characteristics influence people's willingness to delegate decisions to AI. The goal was to replicate the main findings by van den Bergen (2023) while making minor adjustments to rule out some alternative explanations. We compared delegation rates for six different decision conditions, of which one entailed an easy choice and the others were all difficult choices, each with unique characteristics. The hypotheses were tested in two contexts: deciding between two hotels and deciding between different numbers of coins in a gambling game.

Our results suggest that delegation rates are lowest when one alternative is dominant compared to the other, followed by a trade-off condition in which one option displays balanced values and the other option has extreme values (although there is no difference in delegation rates between these first two conditions in the hotel context). Delegation rates are higher for conditions with two options that are equally good or equal but neither good nor bad, as well as for trade-offs between two options that display a mirror image of each other in terms of attribute values. Delegation rates are highest for decisions between two options that are equally bad.

To optimally make use of AI in decision-making situations in real-life, more research is needed into delegation to AI by addressing some limitations of the current study. It is important to take individual differences into account and to understand people's reasons for not trusting and thus not delegating to AI even better.

References

Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI and Society*, *35*(3), 611–623. https://doi.org/10.1007/s00146-019-00931-w

Bettman, J. R., Johnson, E. J., Luce, M. F., & Payne, J. W. (1993). Correlation, Conflict, and Choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 931–951. https://doi.org/10.1037/0278-7393.19.4.931

Bobadilla-Suarez, S., Sunstein, C. R., & Sharot, T. (2017). The intrinsic value of choice: The propensity to under-delegate in the face of potential gains and losses. *Journal of Risk and Uncertainty*, *54*(3), 187–202. https://doi.org/10.1007/s11166-017-9259-x

Bohnet, I., & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior and Organization*, *55*(4 SPEC.ISS.), 467–484. https://doi.org/10.1016/j.jebo.2003.11.004

Braun, R. (2019). Artificial Intelligence: Socio-Political Challenges of Delegating Human Decision-Making to Machines. *IHS Working Paper Series*, *6*. https://ideas.repec.org/p/ihs/ihswps/6.html

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brockman, G., Murati, M., & Welinder, P. (2023). *OpenAI API [Software]*. https://openai.com

Broniarczyk, S. M., & Griffin, J. G. (2014). Decision Difficulty in the Age of Consumer Empowerment. *Journal of Consumer Psychology*, *24*(4), 608–625. https://doi.org/10.1016/j.jcps.2014.05.003

Butler, J. V., & Miller, J. B. (2018). Social risk and the dimensionality of intentions. *Management Science*, *64*(6), 2787–2796. https://doi.org/10.1287/mnsc.2016.2694

Candrian, C., & Scherer, A. (2022). Rise of the machines: Delegating decisions to autonomous AI. *Computers in Human Behavior*, *134*(April), 107308. https://doi.org/10.1016/j.chb.2022.107308

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, *56*(5), 809–825. https://doi.org/10.1177/0022243719851788

Chatterjee, S., & Heath, T. B. (1996). Conflict and loss aversion in multiattribute choice: The effects of trade-off size and reference dependence on decision difficulty. *Organizational Behavior and Human Decision Processes*, *67*(2), 144–155. https://doi.org/10.1006/obhd.1996.0070

Davenport, T. H., & Harris, J. G. (2005). Automated decision making comes of age. *MIT Sloan Management Review*, *46*(4), 83–89.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical Versus Actuarial Judgment Methods of Judgment and Means of Comparison. *Science*, *243*(4899), 1668–1674. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.459.7990&rep=rep1&type=pdf

Dhar, R. (1997). Consumer preference for a no-choice option. *Journal of Consumer Research*, *24*(2), 215–231. https://doi.org/10.1086/209506

Dhar, R., & Simonson, I. (2003). The Effect of Forced Choice on Choice. *Journal of Marketing Research*, *40*(2), 146–160. https://doi.org/10.1509/jmkr.40.2.146.19229

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, *64*(3), 1155–1170. https://doi.org/10.1287/mnsc.2016.2643

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fishburn, P. C. (1968). Utility Theory. *Management Science*, *14*(5), 335–378. https://doi.org/10.1287/mnsc.14.5.335

Frost, R. O., & Shows, D. L. (1993). The nature and measurement of compulsive indecisiveness. *Behaviour Research and Therapy*, *31*(7). https://doi.org/10.1016/0005-7967(93)90121-A

Galanter, E. (1962). The Direct Measurement of Utility and Subjective Probability. *The American Journal of Psychology*, *75*(2), 208–220. https://doi.org/10.2307/1419604

Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective and Behavioral Neuroscience*, *15*(3), 523–536. https://doi.org/10.3758/s13415-015-0347-6

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2022). lab.js: A free, open, online study builder. *Behavior Research Methods*, *54*(2), 556–573. https://doi.org/10.3758/s13428-019-01283-5

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*(6), 995–1006. https://doi.org/10.1037/0022-3514.79.6.995

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kim, J., Novemsky, N., & Dhar, R. (2013). Adding Small Differences Can Increase Similarity and Choice. *Psychological Science*, *24*(2), 225–229. https://doi.org/10.1177/0956797612457388

Leana, C. R. (1986). *Predictors and Consequences of Delegation Author (s): Carrie R . Leana Published by : Academy of Management Stable URL : https://www.jstor.org/stable/255943 REFERENCES Linked references are available on JSTOR for this article : reference # references _ . 29*(4), 754–774.

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data and Society*, *5*(1), 1–16. https://doi.org/10.1177/2053951718756684

Leyer, M., & Schneider, S. (2020). Me, you or ai? How do we feel about delegation. *27th European Conference on Information Systems - Information Systems for a Sharing Society, ECIS 2019*, 0–17.

Lloyd, J., Doll, H., Hawton, K., Dutton, W. H., Geddes, J. R., Goodwin, G. M., & Rogers, R. D. (2010). How psychological symptoms relate to different motivations for gambling: An online study of internet gamblers. *Biological Psychiatry*, *68*(8), 733–740. https://doi.org/10.1016/j.biopsych.2010.03.038

Logg, J. (2018). Theory of Machine: When Do People Rely on Algorithms? *SSRN Electronic Journal*, *17*. https://doi.org/10.2139/ssrn.2941774

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*(April 2018), 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, *46*(4), 629–650.

https://doi.org/10.1093/jcr/ucz013

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2020). Resistance to medical artificial intelligence is an attribute in a compensatory decision process: Response to pezzo and beckstead (2020). *Judgment and Decision Making*, *15*(3), 446–448. https://doi.org/10.1017/s1930297500007233

Louviere, J. J., & Islam, T. (2008). A comparison of importance weights and willingness-to-pay measures derived from choice-based conjoint, constant sum scales and best–worst scaling. *Journal of Business Research*, *61*(9), 903–911. https://doi.org/https://doi.org/10.1016/j.jbusres.2006.11.010

Luce, M. F., Payne, J. W., & Bettman, J. R. (1999). Emotional Trade-Off Difficulty and Choice. *Journal of Marketing Research*, *36*(2), 143–159. https://doi.org/10.1177/002224379903600201

Mahmud, H., Islam, A. K. M. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change*, *175*(November 2021), 121390. https://doi.org/10.1016/j.techfore.2021.121390

Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of "datification." *Journal of Strategic Information Systems*, *24*(1), 3–14. https://doi.org/10.1016/j.jsis.2015.02.001

Owens, B. D., Grossman, Z., & Fackler, R. (2014). *The Control Premium : A Preference for Payoff Autonomy Author (s): David Owens , Zachary Grossman and Ryan Fackler Source : American Economic Journal : Microeconomics , November 2014 , Vol . 6 , No . 4*

Published by : American Economic Association Stable. *6*(4), 138–161.

Pezzo, M. V., & Beckstead, J. W. (2020). Patients prefer artificial intelligence to a human provider, provided the ai is better than the human: A commentary on longoni, bonezzi and morewedge (2019). *Judgment and Decision Making*, *15*(3), 443–445. https://doi.org/10.1017/s1930297500007221

Rassin, E. (2007). A psychological theory of indecisiveness. *Netherlands Journal of Psychology*, *63*(1), 1–11. https://doi.org/10.1007/bf03061056

Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, *37*(3), 409–425. https://doi.org/10.1086/651235

Schmidt, P., Biessmann, F., & Teubner, T. (2020). Transparency and trust in artificial intelligence systems. *Journal of Decision Systems*, *29*(4), 260–278. https://doi.org/10.1080/12460125.2020.1819094

Scholten, M., & Sherman, S. J. (2006). Tradeoffs and theory: The double-mediation model. *Journal of Experimental Psychology: General*, *135*(2), 237–261. https://doi.org/10.1037/0096-3445.135.2.237

Shaddy, F., Fishbach, A., & Simonson, I. (2021). Trade-Offs in Choice. *Annual Review of Psychology*, *72*, 181–206. https://doi.org/10.1146/annurev-psych-072420-125709

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*(1), 11–36. https://doi.org/https://doi.org/10.1016/0010-0277(93)90034-S

Simonson, I., & Tversky, A. (1992). Choice in Context: Tradeoff Contrast and Extremeness Aversion. *Journal of Marketing Research*, *29*(3), 281–295.

StataCorp. (2021). *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC.

Steffel, M., & Williams, E. F. (2018). Delegating decisions: Recruiting others to make choices we might regret. *Journal of Consumer Research*, *44*(5), 1015–1032. https://doi.org/10.1093/jcr/ucx080

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, *211*(4481), 453–458. https://doi.org/10.1126/science.7455683

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics*, *106*(4), 1039–1061. https://doi.org/10.2307/2937956

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*(3).

van den Bergen, D. (2023). *Delegating Subjective Decisions to AI: The Effect of Choice Characteristics*. Eindhoven University of Technology.

von Eschenbach, W. J. (2021). Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philosophy and Technology*, *34*(4), 1607–1622. https://doi.org/10.1007/s13347-021-00477-0

Willemsen, M. C. (2002). *Explaining Asymmetries in Preference Elicitation*. Technische Universiteit Eindhoven.

Willemsen, M. C., Graus, M. P., & Knijnenburg, B. P. (2016). Understanding the role of latent feature Willemsen, M. C., Graus, M. P., & Knijnenburg, B. P. (2016). Understanding the role of latent feature diversification on choice difficulty and satisfaction. User Modeling and User-Adapted Interaction, 26(4), 347–38. *User Modeling and User-Adapted Interaction*, *26*(4), 347–389. https://doi.org/10.1007/s11257-016-9178-6

Willemsen, M. C., & Keren, G. (2002). Negative-based prominence: The role of negative features in matching and choice. *Organizational Behavior and Human Decision Processes*, *88*(2), 643–666. https://doi.org/10.1016/S0749-5978(02)00006-7

Xu, J., Jiang, Z., & Dhar, R. (2013). Mental representation and perceived similarity: How abstract mindset aids choice from large assortments. *Journal of Marketing Research*, *50*(4), 548–559. https://doi.org/10.1509/jmr.10.0390

Yukl, G., & Fu, P. P. (1999). Determinants of Delegation and Consultation by Managers Linked references are available on JSTOR for this article : Determinants of delegation and consultation by managers. *Journal of Organizational Behavior*, *20*(2), 219–232.

Appendices

Appendix A. Sample size justification

To calculate the needed sample size, the results from van den Bergen (2023) were analyzed. As most decision conditions were the same as in the current study and the effect size is expected to be comparable. The mean values and standard deviations of the delegation rate from the previous study are summarized per context and decision condition in Table 21.

Table 21

Delegation Rates per Decision Condition of van den Bergen (2023)

	Hotel context		Coin context	
Decision condition	M	SD	M	SD
Dominant	17.71	30.68	20.14	32.83
Equally good	34.38	37.81	46.87	41.31
Equally bad	35.42	40.88	50	37.30
Trade-off 1	30.21	33.53	21.53	32.65
Trade-off 2	31.77	39.05	59.03	40.71

Note: Mean and standard deviation per decision condition for both the hotel and the coin context of the study of which the current study is a replication

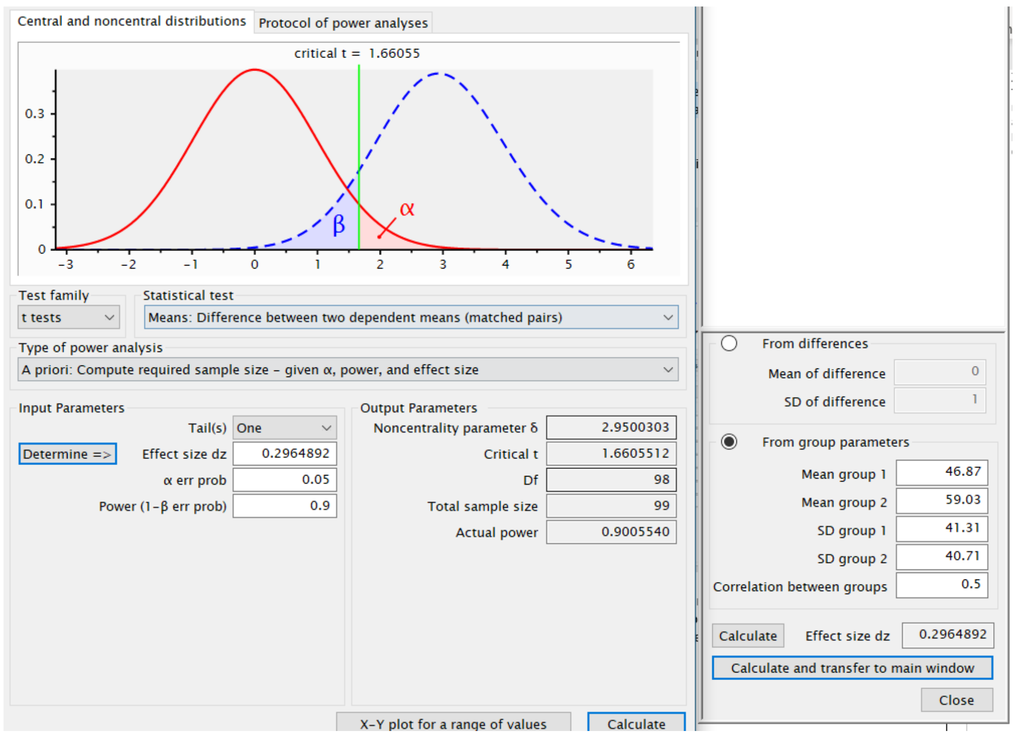
In the current study, the smallest effect size of interest is the one between the equal condition (which, in the previous study, was the equally good condition) and the trade-off 2 condition (in which one option is the mirror image of the other option).

The sample size was calculated in G*Power (t-tests, Means: Difference between two independent means (two groups), a priori) (Faul et al., 2007). The test is one-tailed because we can assume that the results will be comparable to the previous study. We use a power of 0.9 and an alpha level of 0.05.

If we take the mean and SD from the gambling context, the expected effect size is 0.296 and the required sample size would be 99.

Figure 8

*Screenshot of a Priori Sample Size Calculation in G*Power*



Note: The effect size is determined using the mean and standard deviation values from the coin context of van den Bergen (2023). With an alpha level of 0.05 and 0.9 power, the total needed sample size is calculated as 99 participants.

For both contexts, the stimuli values in the current study are adjusted in such a way that they resemble the way the values were presented in the coin context of the previous study. This is why the hypotheses are formulated in such a way that the order of delegation rates between conditions is comparable to the gambling context of the previous study. For this reason, the values from the gamble context are useful and we chose a sample size of 100 participants.

Appendix B. Correlations between control variables

Table 22

Correlation Between Control Variables Hotel Context

	Indecisiveness	Age	DT	Log(DT)	Count trial number	Exchange rate
Indecisiveness						
Age	-.27*					
DT	-.01	.01				
Log(DT)	-.08**	.08**	.53***			
Count trial number	0	0	-.08**	-.26***		
Exchange rate	.06*	-.06*	-.01	-.06*	.01	

Note: DT is the decision time. The exchange rate is the square root of the absolute difference in the calculated exchange rate, based on the point-allocation task.

***p<=.001, **p<=.01, *p<=.05

Table 23

Correlation Between Control Variables Coin Context

	Gambling enjoyment	Indecisiveness	Age	DT	Log(DT)	Count trial number
Gambling enjoyment						
Indecisiveness	.01					
Age	-.19	-.27*				
DT	.015	-.015	.02			
Log(DT)	.17***	-.17***	.17***	.45***		
Count trial number	0	0	0	-.08**	-.34***	

Note: DT is the decision time.

***p<=.001, **p<=.01, *p<=.05

Appendix C. Logistic regression models

Table 24

Full Logistic Regression Model for the Hotel Context

	Odds ratio	Std. err.	P>|z|	[95% conf. interval]	
Dominant	.45	.12	.002	.28	.75
Equally bad	1.71	.41	.024	1.07	2.73
Equally good	.96	.23	.876	.60	1.55
Trade-off 1	.62	.16	.059	.38	1.02
Trade-off 2	.77	.19	.294	.48	1.25
Male	1.97	.85	.117	.84	4.58
No sex	1.49	1.41	.675	.23	9.49
Age	1.02	.02	.674	.97	1.05
Student	.66	.38	.472	.21	2.07
Log(DT)	.02	.02	.000	.00	.15
Log(DT)^2	1.22	.07	.001	1.08	1.37
√Diff. exch. rate	.82	.11	.136	.64	1.06
Indecisiveness	1.66	.55	.129	.86	3.18
Trial count	.96	.01	.006	.93	.99

Note: Logistic regression model with all conditions and control variables included (with equally neutral as reference condition and female as reference gender). Odds ratios, standard errors, p-values and 95% confidence intervals are reported for each of the predictor variables.

Table 25

Part of the Model with Dominant as Reference Condition for the Hotel Context

	Odds ratio	Std. err.	P>|z|	[95% conf. interval]	
Condition					
Equally bad	3.77	.96	.000	2.29	6.20
Equally good	2.12	.55	.004	1.28	3.51
Equally neutral	2.20	.56	.002	1.33	3.63
Trade-off 1	1.37	.36	.231	.82	2.30
Trade-off 2	1.70	.44	.040	1.03	2.83

Note: Logistic regression model with dominant as the reference condition. Control variables were included in the model, but not reported in the table.

Table 26

Full Logistic Regression Model for the Coin Context

	Odds ratio	Std. err.	P>|z|	[95% conf. interval]	
Dominant	.21	.06	.000	.12	.35
Equally bad	2.35	.56	.000	1.47	3.75
Equally good	.78	.19	.308	.49	1.25
Trade-off 1	.42	.10	.000	.26	.68
Trade-off 2	.99	.23	.962	.62	1.57
Male	1.40	.71	.501	.52	3.77
No sex	.42	.50	.463	.04	4.25
Age	1.05	.03	.049	1.00	1.10
Student	.56	.40	.413	.14	2.27
Log(DT)	.66	.09	.001	.51	.85
Indecisiveness	1.36	.53	.428	.63	2.93
Gamble enjoyment	1.23	.47	.597	.58	2.61
Trial count	1.00	.01	.766	.96	1.02

Note: Logistic regression model with all conditions and control variables included (with equally neutral as reference condition and female as reference gender). Odds ratios, standard errors, p-values and 95% confidence intervals are reported for each of the predictor variables.

Table 27

Part of the Model with Dominant as Reference Condition for the Coin Context

	Odds ratio	Std. err.	P>|z|	[95% conf. interval]	
Equally bad	11.28	3.09	.000	6.60	19.28
Equally good	3.76	1.01	.000	2.23	6.37
Equally neutral	4.80	1.28	.000	2.85	8.11
Trade-off 1	2.02	.55	.010	1.18	3.43
Trade-off 2	4.75	1.27	.000	2.82	8.01

Note: Logistic regression model with dominant as the reference condition. Control variables were included in the model, but not reported in the table.

Table 28

Full Logistic Regression Model for Both Contexts Combined

	Odds ratio	Std. err.	P>|z|	[95% conf. interval]	
Dominant	.32	.06	.000	.23	.46
Equally bad	1.92	.31	.000	1.40	2.62
Equally good	.85	.14	.335	.62	1.18
Trade-off 1	.51	.08	.000	.36	.70
Trade-off 2	.87	.14	.376	.63	1.19
Male	1.74	.69	.162	.80	3.79
No sex	.54	.48	.492	.09	3.12
Age	1.02	.02	.315	.98	1.06
Student	.59	.34	.357	.19	1.80
Log(DT)	.00	.00	.000	.00	.00
Log(DT)^2	1.46	.07	.000	1.32	1.60
Indecisiveness	1.63	.50	.112	.89	2.97
Gamble enjoyment	1.01	.31	.966	.56	1.84

Note: Logistic regression model with all conditions and control variables included (with equally neutral as reference condition and female as reference gender). Odds ratios, standard errors, p-values and 95% confidence intervals are reported for each of the predictor variables.

Table 29

Part of the Model with Dominant as Reference Condition for Both Contexts Combined

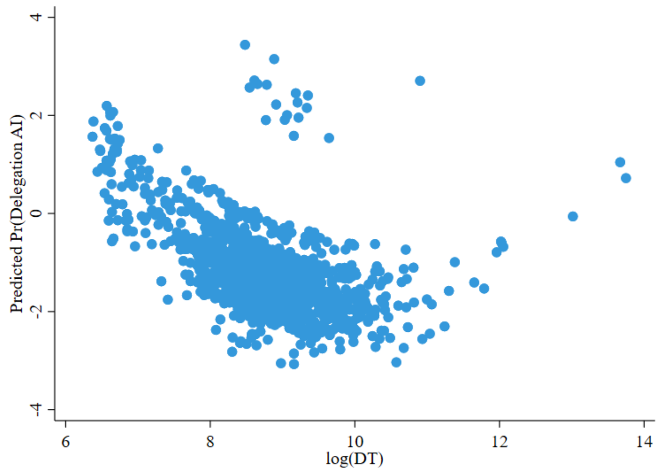
	Odds ratio	Std. err.	P>|z|	[95% conf. interval]	
Equally bad	5.93	1.05	.000	4.19	8.40
Equally good	2.67	.48	.000	1.89	3.79
Equally neutral	3.09	.55	.000	2.19	4.37
Trade-off 1	1.59	.29	.010	1.12	2.27
Trade-off 2	2.70	.48	.000	1.90	3.82

Note: Logistic regression model with dominant as the reference condition. Control variables were included in the model, but not reported in the table.

Appendix D. The effect of decision time on delegation

Figure 9

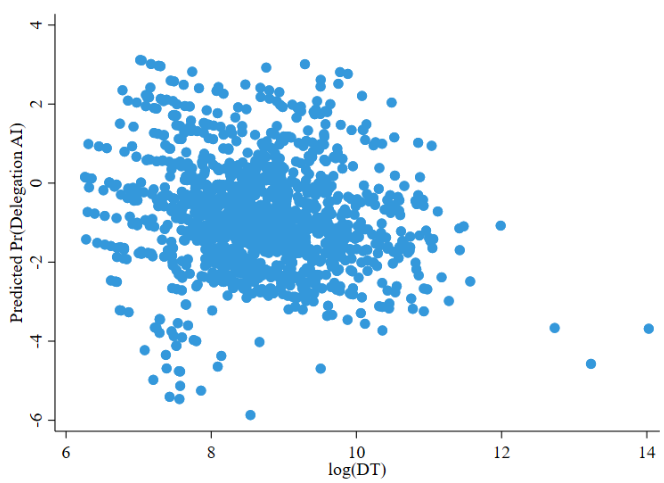
The Relation Between Delegation and Decision Time in the Hotel Context



Note: The predicted proportion that is delegated to AI is plotted against the log-transformed decision time

Figure 10

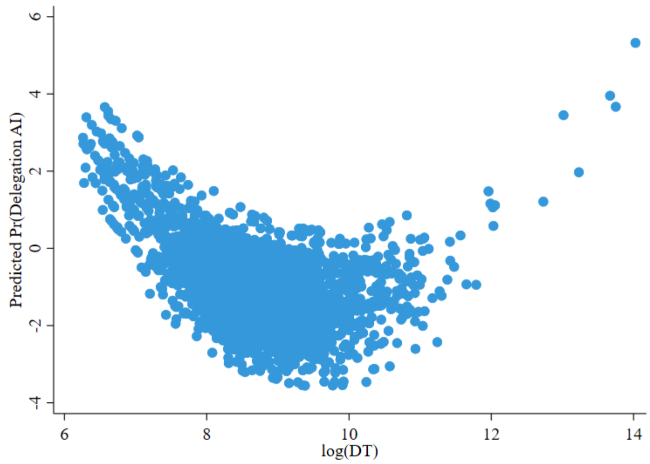
The Relation Between Delegation and Decision Time in the Coin Context



Note: The predicted proportion that is delegated to AI is plotted against the log-transformed decision time

Figure 11

The Relation Between Delegation and Decision Time in Both Contexts Combined

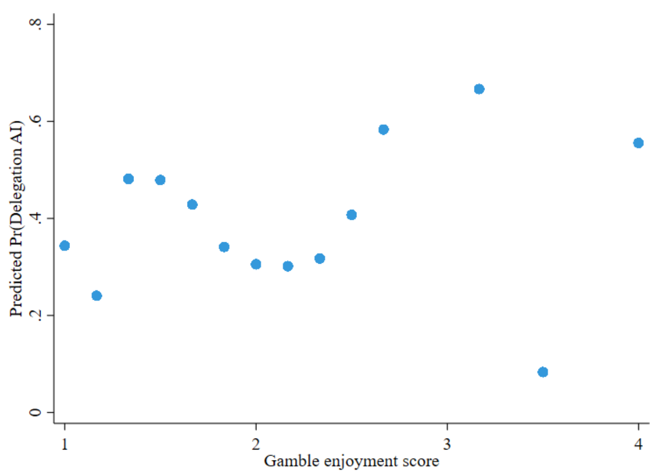


Note: The predicted proportion that is delegated to AI is plotted against the log-transformed decision time

Appendix E. The effect of gambling enjoyment on delegation

Figure 12

The Relation Between Delegation and Gambling Enjoyment Score in the Coin Context



Note: The predicted proportion that is delegated to AI is plotted against the gamble enjoyment score.