Preferences for ChatGPT vs. Human Advisers: Do the Adviser and Cognitive Load Predict Risky Decision-Making?

Joep G. H. Voorter – s1098580

Research Project 3: Bachelor's Thesis (SOW-PSB3RS61E)
28th of February, 2026
Bachelor Psychology

Prof. A. G. Sanfey and Dr. C. E. Rățală
Faculty of Social Sciences, Radboud University

Table of Contents

Abstract: Participants aged 18–65 years old were recruited to participate in the Trust Game, which was an online simplified version of the one used in Klingbeil et al. (2024). In half of the trials, they received advice from ChatGPT, and in the other half, it was from another participant. The first block did not induce cognitive load, and in the second block cognitive load was induced by including time constraints. Across all four conditions, the number of times they took the given advice was measured. No significant effect on rates of advice taking was found for source of advice, cognitive load, or for their interaction. These results did not replicate most earlier findings. It is possible that factors like the way we manipulated cognitive load or the adjustments to the experimental design, as well as the research sample, account for these differences.

Keywords: decision-making, social risk, advice-taking, following advice, ChatGPT, cognitive load, time pressure, financial decisions, trusting behaviour, algorithm aversion, algorithm trust

Decision-making constitutes a vast proportion of human behaviour. When the associated outcomes are uncertain, we speak of *risky decision-making* (Birnbaum, 2008; Gafni et al., 2013; Johnson & Busemeyer, 2010; Starmer, 2000). People have different ways to go about it when they encounter a risky decision (Dewberry et al., 2013; Thunholm, 2004), and as the available information is typically limited, the individual can compensate for this by enquiring for information and advice from outside sources (Bonaccio & Dalal, 2006; Yaniv, 2004). The way they utilise the advice they receive depends on a number of factors, such as the source of advice (SoA)—who created the advice—as well as the perception of the source's level of expertise, the closeness and trust levels of the source, and how much they identify with the source (Feng & MacGeorge, 2006; Mesbah et al., 2021; Meshi et al., 2012; Van Swol & Sniezek, 2005). These factors all relate to the behavioural response, or *risky decision-making behaviour*. They even extend to another characteristic of the adviser, namely whether it is a non-sentient artificial intelligence (AI) adviser or another human being (Chua et al., 2023; Hertz & Wiese, 2019; Klingbeil et al., 2024; Logg et al., 2019; Osborne & Bailey, 2025; Pearson et al., 2019; Yang et al., 2025).

Over the last couple of years, AI has been becoming increasingly integrated into our society (Rahwan et al., 2019) and has resulted in a variety of attitudes. One study showed that for conditions where either the AI or the human adviser's perceived level of expertise was high, participants tended to choose the high-expertise adviser; whether this was a human or AI had no effect. However, for conditions where perceived expertise was high for both the AI and human adviser, the participants preferred the human adviser (Pearson et al., 2019). Another study confirmed the findings that observed preferences for expert advisors, as their judges displayed similar behaviour; they also showed judges preferred AI advice when this was in line with their own assessment and consulted AI and human sources equally for the other trials (Mesbah et al., 2021). Research examining the role of the field or topic for which the advice is given showed that for advice on technology and software or analytical tasks, participants were more prone to favour AI advice; for areas such as relationships or social tasks, they displayed a preference for human advice; and human advisers were also preferred when the type of task was not clear yet (Hertz & Wiese, 2019; Osborne & Bailey, 2025). Another study found a positive association between attitude towards AI and advice taking (AT) from AI, but for decisions with high risk levels there was no effect from attitude, and participants showed stronger preferences for AI advice for riskier decisions (Chua et al., 2023). Another study found consistent preferences for AI advice in all of their conditions when their participants received advice on tasks where they had to make a variety of estimations and predictions (Logg et al., 2019).

Another element that is often linked with risky decision-making is cognitive load (CL). Research has shown that higher levels of CL tend to have a negative effect on financial decision-making, but also that high levels of CL when making financial decisions led participants to show increased risk-taking behaviour, as well as placing higher bids (Benjamin et al., 2013; Blaywais & Rosenboim, 2019; Lahav et al., 2025). High CL was also found to increase impulsivity when making financial decisions, in the sense that participants often chose immediate low-value rewards over delayed high-value rewards (Hinson et al., 2003). Another study also measured increased impulsiveness when CL was increased, as well as lower levels of trust (Samson & Kostyszyn, 2015). Reductions in risk-taking behaviour for financial decisions as a consequence of increasing CL—which is contradictory to some other findings—were also found in another set of experiments by Deck and Jahedi (2015).

As the studies on how CL affects risky decision-making show contradictory findings, and though various studies have indicated which SoA (AI or human) is preferred among various topics, there is very little research on the effect of SoA on risky decision-making and how CL can play a moderating role. The effect of SoA is closely related to the concepts of *algorithm aversion* (Dietvorst et al., 2015) and *algorithm appreciation* (Logg et al., 2019), which refer to a general attitude towards AI but do not necessarily predict attitude-conforming behaviour where AI is involved, such as AT. Though a lot of research has shown how AI interventions have an alleviating effect on CL (Chen & Chang, 2024; Feng, 2025; Hong & Guo, 2025; Hor et al., 2025), research hardly covers the opposite causal direction of how AI-related outcomes are affected by CL, nor does it often cover the interaction of CL and SoA and how this affects risky decision-making.

Our research question encompasses situations where the *source of risk* is social—meaning the outcome of the decision depends on another human, rather than on chance—and narrows risky decision-making down to the AT aspect. We will focus on the effect of SoA on AT, whether this effect is moderated by CL, and whether there is an association between CL and AT. We hypothesise that SoA has an effect on AT (H_1), there is an association between CL and AT (H_2), and the effect of SoA on AT is moderated by CL (H_3).

Gaining insight into the way these three variables relate to each other benefits society, as decision-making and CL directly relate to sustainable development goals: *no poverty* (1), *zero hunger* (2), and *good health and well-being* (3). Research shows that cognitive functions, such as inhibitory control and planning, are essential in making optimal economic decisions. The various ways in which poverty has a detrimental effect on cognitive functioning can increase the individual's CL. When this results in an inability to make well-informed financial decisions, it may create a vicious cycle, where these negative effects perpetuate poverty (Dean et al., 2018; Mani et al., 2013). Further research suggests that scarcity of resources increases CL for individuals and that their engagement with their problems takes up most of their attentional resources, which potentially impairs their ability to make well-informed optimal financial decisions (Shah et al., 2012).

Our study also benefits our society by providing insight into the possible risk of overreliance on AI advice. For a lot of situations, AI-generated advice may prove to be highly accurate, and its applications can be very effective (Rao et al., 2023). However, there is sufficient reason for concern about practices where excessively relying on AI-generated info leads to detrimental repercussions (Araujo et al., 2022; Klingbeil et al., 2024; Logg et al., 2019). In a variety of fields, such as decision-making processes for psychological treatment and delegating managerial decisions, individuals display an excessive proclivity for trusting algorithmic recommendations (Candrian & Scherer, 2022; Jacobs et al., 2021).

The study from Klingbeil et al. (2024) investigated the question of whether SoA has an effect on AT and whether participants would display overreliance towards ChatGPT-generated advice. They found that participants indeed would conform to the ChatGPT advice more frequently compared to the human advice; that they would even conform to ChatGPT advice when this advice was against their own judgement; and, more remarkably, that they would even prefer ChatGPT advice over expert advice. For their design, the source of risk was social as well, and players had to participate in a financial decision task where their decisions were based on trust and the choices of another player they competed with. We will use a simplified version of the same financial decision task where the interaction is replaced by a summary of the opponent's last three moves. Our version has one block without CL and another block with high CL—representing the two conditions of our independent variable (IV) CL—where both blocks contain eight consecutive trials with ChatGPT advice and eight consecutive trials with advice from another participant—representing the two conditions of SoA, which is our other IV. The selected choice is compared to the given advice, which is how our dependent variable (DV) AT will be measured. We predict that AT rates will be significantly different among all four conditions.

Methods

Participants

We conducted an a priori power analysis for a two-tailed logistic regression using the G*power 3.1.9.7 program (Faul et al., 2009) and performed subsequent interim analyses for adjusting the sample size according to the number of measurements per participant (see Appendix A for the exact procedure). To yield the target statistical power ($1 - \beta = .80$), having $p_1 = .55$ (Faul et al., 2009) and $OR = 1.5$ as effect size expectations, we need at least 104 participants in order for these effects to be statistically significant ($\alpha = .05$).

We collected data from 179 participants, of whom 141 participants completed the experiment, yielding usable data. The participants in the social source of risk condition constitute our research sample ($N = 71$). Concerning gender and age distributions, we recruited participants aged 18–65 years of all genders. We recruited participants by means of our social circles and group chats, as well as through the SONA platform, where BSc Psychology students from Radboud University receive 0.5 SONA credit as reimbursement.

As shown in Table B1, $n_{\text{women}} = 64$ female and $n_{\text{men}} = 7$, and the data from Table B2 shows that almost all our participants ($n = 69$) were aged 18–24 ($M = 19.9$, $SD = 3.9$), and 66.2% of the sample were aged 18–19 ($n_{\text{age<20}} = 47$). Out of 179 participants, 134 participants completed the study via SONA and thus are BSc Psychology students at Radboud University. We adhered to the customary ethical procedure, so this study was carried out in accordance with the guidelines of the ECSS of the RU, which are outlined in protocol ECSW-2025-109.

Materials

We measured risk-taking behaviour using a *decision task*, which, for the social version we used, was derived from the Trust Game with the structure that was first introduced by Berg et al. (1995). A validity analysis conducted by Banerjee et al. (2021) shows that the Trust Game's external validity is high and that it correlates strongly with other scales measuring social trust. The decision task was divided into two blocks with 16 trials in each block. For each trial, the participant was presented with the other participant's *history*, and the participant had to decide whether to respond with IN or OUT while advice was presented (see Appendix C for the specific instructions and elaboration on the games' rules). Our design resembles the operationalisation from Klingbeil et al. (2024), but it is much more simplified due to our research restraints (see Appendix D for the design we used). We measured AT by means of the decision of the participant.

CL, which was one of our IVs, was operationalised by adding time pressure for the high-load condition. This method of manipulating CL was the most suitable fit with the rest of our design, and imposing time constraints has been successfully used for manipulating CL in previous research (Deck et al., 2017; Howarth & Anderson, 2007; Klapproth, 2008; Orthey et al., 2019). CL was a binary variable where the low-load condition was measured during the first block of 16 trials without time constraints (see Figure D1 for design), and the high-load condition was measured during the second block of 16 trials where *time pressure* was induced by displaying a timer (see Figure D2 for design). CL was a within-subject variable that was not randomised to prevent spillover effects.

SoA, being the other IV, was binary as well with ChatGPT and human conditions, and this was a within-subject factor that was randomised within each block of 16 trials. For each condition, with a duration of eight consecutive trials, the advice was presented as though another participant with a code—for the human condition—or ChatGPT—for the ChatGPT condition—had provided it (see Appendix D for presentations). In fact, participants were only led to believe advice came from those sources. This was done by telling the participant they would be competing against another player who is also doing the decision task, by displaying a different participant code each round for the other players, and by asking the participant to give advice to another participant during the game.

Procedure

The experiment was designed in alignment with the levels of the methodological framework that apply to this project (see Appendix E). When entering the Qualtrics survey, the participants were presented with an introduction and information letter (see Appendix F); on the next page, they had to read the consent form (see Appendix G). They had to answer two questions: "I give explicit consent for processing the following personal data about me: age and gender," with answer options *yes* and *no*; and "Do you agree to participate in this study?" with answer options "Yes, I agree to participate in this study," and "No, I do not want to participate in this study." If they answered both questions affirmatively, they would proceed to the next two questions: "Please indicate your gender" with answer options *female, male, other,* and *prefer not to say*; and "Please indicate your age," where they responded by using a slider with an 18–65 range. Participants were then randomly divided into the social or non-social group (source of risk was a between-subject variable for fellow researchers).

The participants from the social group entered the decision task and had to read the task instructions; as a comprehension check, they had to answer three questions with answer options *€0, €10, €20,* and *€40* correctly to proceed to the trials (see Appendix C for task instructions and comprehension questions). Each trial displayed the history, after 2 s the advice, and after 2 more s participant could select a button. Between every trial, the inter-trial screen was shown for about 2 s (see Figure D2). The first block started with six baseline trials, followed by six trials of giving advice to another participant, based on their history (see Figure D1). Then eight trials with human or ChatGPT advice and eight subsequent trials from the other SoA condition followed. After the first block the warning time constraint screen was displayed (see Figure D2). The second block began automatically with six trials without advice and again eight trials for both SoA conditions, but with a timer (see Figure D2).

Upon completing the decision task, the participants were redirected to some questions about the choices they made. They had to give two answers to the question "How useful was the advice you got from…" for both *other participants* as well as *ChatGPT* using two sliders, whose scores ranged from 0 to 100. Above the sliders, five level indicators that did not represent any score levels in particular – *not at all useful, slightly useful, moderately useful, very useful*, and *extremely useful* – were displayed in this order from left to right. They had to fill out two questionnaires on neuroticism and self-efficacy (only relevant for fellow researchers). Lastly, the closure and written debriefing—that disclosed the deceptions that were used—were shown to the participants (see Appendix H). As this information was previously unknown to the participants, they were given the option to withdraw their participation consent from the study by sending an e-mail.

Data Analysis

Using IBM SPSS Statistics version 29.0.1.0, preprocessing was conducted by transforming comma-separated values into variables and removing the baseline trials without advice so that 32 cases (representing trials) for each participant remained. It also included the variable column "followed_advice" that displayed the output for AT as a quantitative variable with the levels 0 (where participant decision ≠ given advice) and 1 (where participant decision

= given advice). We changed this variable from numeric to string with value labels *no* (0) and *yes* (1) and changed the role of this variable from "input" to "target", as our DV was a qualitative binary variable. The data in the column 'cognitive_load' displayed CL, which was a qualitative binary variable with the levels *low load* (1) and *high load* (2). The data in the column "advice_condition" displayed SoA, which was a qualitative binary variable with the levels *AI* (2) for ChatGPT and *Human* (3).

We selected all cases for which the source of risk was social—the column "risk_domain" = 2, as this constituted our research sample. Using IBM SPSS Statistics version 29.0.1.0, we conducted a generalised linear mixed model (GLMM) to investigate whether SoA and CL have an impact on AT, which was our binary DV (0, 1). The probability distribution of the analysis was binomial, and it had a logit link function. For fixed effects, we selected SoA, which was our binary IV (human, ChatGPT), and CL, which was also a binary IV (low-load, high-load), as well as the interaction effect (contrasted against CL) of these two IVs. Random intercepts were included for participants' IDs to account for repeated measurements. As all our variables were dichotomous and we used a logit link function, we did not need to apply any centering to our IVs manually.

Results

All of the descriptive statistics for the key variables were almost identical across all nine conditions. The means ranged from .406 to .424; the standard errors ranged from .413 to .418; for the confidence intervals from all conditions, the lower limits ranged from .023 to .025, and the upper limits ranged from .952 to .955. The exact statistics for each variable can be found in Table I1. The correlations between the fixed coefficients of these three key predictors are $r = .498$ between CL and SoA, $r = –.707$ between SoA and the interaction, and $r = –.706$ between CL and the interaction. For all condition sets, we created frequency tables for the number of trials where advice was followed (see Table J1), as well as a frequency table that comprised all 32 trials (see Table J2).

The GLMM analysis showed that the main effect of SoA on AT, which represented H_1, was not significant, $F(1, 2268) = 0.36$, $p = .551$. This suggests participants displayed no significant differences in AT, whether advice was provided by a human or by ChatGPT. The main effect of CL on AT, which represented H_2, was also not significant, $F(1, 2268) = 0.00$, $p = .999$. This suggests that for the low CL trials, AT rates were similar to those from the high CL trials. The interaction effect between CL and SoA on AT, which represented H_3, was not significant either, $F(1, 2268) = 0.07$, $p = .798$. This suggests that the effects of SoA on AT were similar between the low-load and high-load conditions. The random effect for participant ID ($\sigma^2 = 0.05$, $SD = 0.03$, $Z = 1.61$, $p = .107$) was also not significant, so it could not account for unexplained variability within our measurements. The coefficients for the predictors, as well as all other relevant statistics for fixed and random effects, are presented in the overview in Table 1.

Table 1

GLMM Analysis Results: Coefficients for Fixed Effects and Covariance for Random Effects

Effect	Estimate	*SE*	$\chi^2(1)$	*p*	*OR*	95% CI	
						LL	*UL*
Fixed effects							
Intercept	–0.31	1.71	0.03	.858	0.74	0.03	21.14
SoA	–0.07	0.12	0.37	.547	0.93	0.73	1.18
CL	–0.02	0.12	0.03	.857	0.98	0.77	1.24
CL × SoA	0.04	0.17	0.07	.798	1.05	0.75	1.46
Random effects							
Intercept	2.92						
Participants	0.05	0.03		.107		0.01	0.17

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit; SoA = source of advice; CL = cognitive load

Discussion

For this study, we focused on the research question of whether SoA has an effect on AT and whether this effect is moderated by CL and whether there is an association between CL and AT. As the main effects of SoA and CL, as well as their interaction effect, were not significant, all three H_0's are retained, and we cannot conclude that there are associations between SoA, CL, and AT.

Our findings show some overlap with the findings of a study investigating whether judges preferred AI or human advice after making their own estimation. They found that AI and human advice were utilised equally and that other characteristics, such as adviser expertise, had a stronger effect on which SoA was preferred (Mesbah et al., 2021).

Another study where human and AI advice was presented simultaneously also found that perceived expertise determined which advice was being utilised, with one exception being that in a scenario where both advisers' perceived expertise was high, a slight preference for human advice was observed (Pearson et al., 2019). Compared to Yang et al. (2025), our study found conceptually similar results in the sense that CL did not have an effect on algorithm trust, measured as $AT_{ChatGPT} - AT_{human}$. However, it practically contradicted its findings, as Yang et al. (2025) found that time pressure had a positive effect on algorithm trust and AT from an AI, and we manipulated CL by using time pressure. Some of the previous studies that were only loosely related to ours—as they focused on risk-taking behaviour instead of AT—contradicted with our findings in the sense that they found a positive effect on economic risk-taking (Benjamin et al., 2013; Blaywais & Rosenboim, 2019; Lahav et al., 2025) or a negative effect (Deck & Jahedi, 2015; Samson & Kostyszyn, 2015), and we did not find any effect.

The absence of effect from CL can be explained by the fact that various findings of CL use different manipulations, like having participants listen to unsettling noises (Samson & Kostyszyn, 2015) or using a variety of working memory (WM) manipulations (Hinson et al., 2003). But also more specific WM manipulations like sequence memorisation (Ball et al., 2023; Benjamin et al., 2013; Blaywais & Rosenboim, 2019; Deck & Jahedi, 2015; Duffy & Smith, 2014; Lahav et al., 2025; Samson & Kostyszyn, 2015; Whitney et al., 2008; Yang et al., 2025) or pattern memorisation (Gerhardt et al., 2016). As the article by Yang et al. (2025) describes, different facets of what is commonly referred to as CL can have different effects, and factually speaking, in this study we looked at the effects of time pressure. Deck et al. (2017) compared different ways of manipulating cognitive load as well, which would suggest that CL is not one universal concept with a clear general definition.

Our findings also contradicted the findings from Klingbeil et al. (2024), who used the same decision task to show an overreliance on ChatGPT advice. A major difference that could account for this is that Klingbeil et al. (2024) used human experts for human advice and used ChatGPT-generated text, including the reasoning behind the ChatGPT advice, which can have a significant impact. For our experiment, both human and ChatGPT advisers may have ‘felt’ similar to participants as they were doing all of it behind their screen. Another factor that could account for these contradicted findings is that our participants were led to believe they received advice from another player who was, just like them, taking this experiment, making it likely that participants perceived the other player’s level of expertise as similar to their own. If they did, the perceived expertise of our human advisers was considerably lower than the expert advisers from Klingbeil et al. (2024).

Our research design was constrained by some limitations that could possibly explain discrepancies with previous research. As we could not present the human adviser as an expert, and as the claim was made that ChatGPT had been trained on previous data. Perceived expertise of the adviser could have been a confounding variable, especially since a number of studies demonstrated the superior effect of perceived expertise on AT (Feng & MacGeorge, 2006; Mesbah et al., 2021; Meshi et al., 2012; Pearson et al., 2019; Schrah et al., 2006; Van Swol & Sniezek, 2005). Furthermore, from the 179 participants, 134 participants were reimbursed with SONA credits. So the students from BSc Psychology at Radboud constituted the majority of both our social research sample and the non-social research sample. Furthermore, previous findings show that students are prone to display less trustworthy behaviour in these decision tasks (Johnson & Mislin, 2011), which might have affected our results. Because of this, and combined with the fact that 90.1% of our sample was female ($n_{\text{women}} = 64$) and 66.2% was aged below 20 ($n_{\text{age<20}} = 47$), our research has low external validity. Another limitation came with the methodological simplification that participants were only exposed to one stage of the Trust Game. They only had the knowledge of how the game worked and did not experience betrayal or cooperation from an opponent personally. The history they got to see might not have had any personal relevance to them, which might have impacted risky decision-making. So also here our design did not resemble Klingbeil et al. (2024). With our design, participants might feel increasing desensitisation to the significance of their decision, and the *perceived risk* they are taking might decrease over time.

Suggestions for future research with a similar design are to include an extra questionnaire containing statements like the one I made in the previous point, aimed at measuring their subjective experience of 'investment in the game' and whether they really did feel like they were 'making an exciting gamble' every round. In the operationalisation of SoA, it is advisable to control for the perceived expertise of the advisers and to include an implicit and explicit measurement of attitude towards AI, or even ChatGPT in particular. For investigating the effect of CL, specifying CL sufficiently is recommended, for it enables differentiation between the effects of different facets of CL, potentially resulting in a research database that has different subdivisions within the umbrella term 'CL'.

In conclusion, our research could not replicate many of the earlier findings that show a preference for either human or ChatGPT recommendations. Characteristics of our research restraints have resulted in various suggestions for future research directions. As many characteristics of the adviser are embedded in risky decision-making, and reliance on AI algorithms depends on many different factors but yet plays an increasingly crucial role in our modern society, there is a high need for more thorough and concrete research—among a variety of populations—in this field with increasing complexity.

References

Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2022). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & Society, 35*, 611–623. https://doi.org/10.1007/s00146-019-00931-w

Ball, S., Katz, B., Li, F., & Smith, A. (2023). The effect of cognitive load on economic decision-making: A replication attempt. *Journal of Economic Behavior and Organization, 210*, 226–242. https://doi.org/10.1016/j.jebo.2023.03.018

Banerjee, S., Galizzi, M. M., & Hortala-Vallve, R. (2021). Trusting the trust game: An external validity analysis with a UK representative sample. *Games, 12*(3), Article 66. https://doi.org/10.3390/g12030066

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is 'behavioral'? Cognitive ability and anomalous preferences. *Journal of the European Economic Association, 11*(6), 1231–1255. https://doi.org/10.1111/jeea.12055

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122–142. https://doi.org/10.1006/game.1995.1027

Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review, 115*(2), 463–501. https://doi.org/10.1037/0033-295X.115.2.463

Blaywais, R., & Rosenboim, M. (2019). The effect of cognitive load on economic decisions. *Managerial and Decision Economics, 40*(8), 993–999. https://doi.org/10.1002/mde.3085

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes, 101*(2), 127–151. https://doi.org/10.1016/j.obhdp.2006.07.001

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

Chen, C.-H., & Chang, C.-L. (2024). Effectiveness of AI-assisted game-based learning on science learning outcomes, intrinsic motivation, cognitive load, and learning behavior. *Education and Information Technologies, 29*, 18621–18642. https://doi.org/10.1007/s10639-024-12553-x

Chua, A. Y. K., Pal, A., & Banerjee, S. (2023). AI-enabled investment advice: Will users buy it? *Computers in Human Behavior, 138*, Article 107481. https://doi.org/10.1016/j.chb.2022.107481

Dean, E. B., Schilbach, F., & Schofield, H. (2018). Poverty and cognitive function. In C. B. Barrett, M. Carter, J.-P. Chavas, & M. R. Carter (Eds.), *The economics of poverty traps* (pp. 57–118). University of Chicago Press. https://doi.org/10.7208/chicago/9780226574448.003.0003

Deck, C., & Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review, 78*, 97–119. https://doi.org/10.1016/j.euroecorev.2015.05.004

Deck, C., Jahedi, S., & Sheremeta, R. (2017). The effects of different cognitive manipulations on decision making. *Chapman University, Economic Science Institute, Working Paper Series*. https://tinyurl.com/42vh8e5y

Dewberry, C., Juanchich, M., & Narendran, S. (2013). The latent structure of decision styles. *Personality and Individual Differences, 54*(5), 566–571. https://doi.org/10.1016/j.paid.2012.11.002

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

Duffy, S., & Smith, J. (2014). Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *Journal of Behavioral and Experimental Economics, 51*, 47–56. https://doi.org/10.1016/j.socec.2014.01.006

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Feng, B., & MacGeorge, E. L. (2006). Predicting receptiveness to advice: Characteristics of the problem, the advice-giver, and the recipient. *Southern Communication Journal, 71*(1), 67–85. https://doi.org/10.1080/10417940500503548

Feng, L. (2025). Investigating the effects of artificial intelligence-assisted language learning strategies on cognitive load and learning outcomes: A comparative study. *Journal of Educational Computing Research, 62*(8), 1741–1774. https://doi.org/10.1177/07356331241268349

Gafni, A., Walter, S., & Birch, S. (2013). Uncertainty and the decision maker: Assessing and managing the risk of undesirable outcomes. *Health Economics, 22*(11), 1287–1294. https://doi.org/10.1002/hec.2883

Gerhardt, H., Biele, G. P., Heekeren, H. R., & Uhlig, H. (2016). Cognitive load increases risk aversion. *Sonderforschungsbereich 649: Ökonomisches Risiko, 2016*, Article 11. https://doi.org/10.18452/4622

Hertz, N., & Wiese, E. (2019). Good advice is beyond all price, but what if it comes from a machine? *Journal of Experimental Psychology: Applied, 25*(3), 386–395. https://doi.org/10.1037/xap0000205

Hinson, J. M., Jameson, T. L., & Whitney, P. (2003). Impulsive decision making and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(2), 298–306. https://doi.org/10.1037/0278-7393.29.2.298

Hong, X., & Guo, L. (2025). Effects of AI-enhanced multi-display language teaching systems on learning motivation, cognitive load management, and learner autonomy. *Education and Information Technologies, 30*, 17155–17189. https://doi.org/10.1007/s10639-025-13472-1

Hor, T. S. C., Fong, L., Wynne, K., & Verhoeven, B. (2025). Easing the cognitive load of general practitioners: AI design principles for future-ready healthcare. *Technovation, 142*, Article 103208. https://doi.org/10.1016/j.technovation.2025.103208

Howarth, B., & Anderson, A. H. (2007). Introducing objects in spoken dialogue: The influence of conversational setting and cognitive load on the articulation and use of referring expressions. *Language and Cognitive Processes, 22*(2), 272–296. https://doi.org/10.1080/01690960600632796

Jacobs, M., Pradier, M. F., McCoy Jr., T. H., Perlis, R. H., Doshi-Velez, F., & Gajos, K. Z. (2021). How machine-learning recommendations influence clinician treatment selections: The example of antidepressant selection. *Translational Psychiatry, 11*(108). https://doi.org/10.1038/s41398-021-01224-x

Johnson, J. G., & Busemeyer, J. R. (2010). Decision making under risk and uncertainty. *WIREs Cognitive Science, 1*(5), 736–749. https://doi.org/10.1002/wcs.76

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology, 32*(5), 865–889. https://doi.org/10.1016/j.joep.2011.05.007

Klapproth, F. (2008). Time and decision making in humans. *Cognitive, Affective, & Behavioral Neuroscience, 8*(4), 509–524. https://doi.org/10.3758/cabn.8.4.509

Klingbeil, A., Grützner, C., & Schreck, P. (2024). Trust and reliance on AI – An experimental study on the extent and costs of overreliance on AI. *Computers in Human Behavior, 160*, Article 108352. https://doi.org/10.1016/j.chb.2024.108352

Lahav, E., Manos, R., Kashy-Rosenbaum, G., & Sitbon, N. (2025). Trying to think: An experimental study of the impact of cognitive load on financial risk taking by groups. *Finance Research Letters, 74*, Article 106823. https://doi.org/10.1016/j.frl.2025.106823

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*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science, 341*(6149), 976–980. https://doi.org/10.1126/science.1238041

Mesbah, N., Tauchert, C., & Buxmann, P. (2021). Whose advice counts more – man or machine? An experimental investigation of AI-based advice utilization. *Proceedings of the 54th Hawaii International Conference on System Sciences, 8*(496), 4083–4092. https://doi.org/10.24251/HICSS.2021.496

Meshi, D., Biele, G., Korn, C. W., & Heekeren, H. R. (2012). How expert advice influences decision making. *PLoS ONE, 7*(11), Article e49748. https://doi.org/10.1371/journal.pone.0049748

Orthey, R., Palena, N., Vrij, A., Meijer, E., Leal, S., Blank, H., & Caso, L. (2019). Effects of time pressure on strategy selection and strategy execution in forced choice tests. *Applied Cognitive Psychology, 33*(5), 974–979. https://doi.org/10.1002/acp.3592

Osborne, M. R., & Bailey, E. R. (2025). Me vs. the machine? Subjective evaluations of human- and AI-generated advice. *Scientific Reports, 15*, Article 3980. https://doi.org/10.1038/s41598-025-86623-6

Pearson, C. J., Geden, M., & Mayhorn, C. B. (2019). Who's the real expert here? Pedigree's unique bias on trust between human and automated advisers. *Applied Ergonomics, 81*, Article 102907. https://doi.org/10.1016/j.apergo.2019.102907

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A., . . . Wellman, M. (2019). Machine behaviour. *Nature, 568*, 477–486. https://doi.org/10.1038/s41586-019-1138-y

Rao, A., Pang, M., Kim, J., Kamineni, M., Lie, W., Prasad, A. K., Landman, A., Dreyer, K., & Succi, M. D. (2023). Assessing the utility of ChatGPT throughout the entire clinical workflow: Development and usability study. *Journal of Medical Internet Research, 25*, Article e48659. https://doi.org/10.2196/48659

Rutterford, C., Copas, A., & Eldridge, S. (2015). Methods for sample size determination in cluster randomized trials. *International Journal of Epidemiology, 44*(3), 1051–1067. https://doi.org/10.1093/ije/dyv113

Samson, K., & Kostyszyn, P. (2015). Effects of cognitive load on trusting behavior – An experiment using the trust game. *PLoS ONE, 10*(5), Article e0127680. https://doi.org/10.1371/journal.pone.0127680

Schrah, G. E., Dalal, R. S., & Sniezek, J. A. (2006). No decision-maker is an island: Integrating expert advice with information acquisition. *Journal of Behavioral Decision Making, 19*(1), 43–60. https://doi.org/10.1002/bdm.514

Shah, A. K., Mullainathan, S., & Shafir, E. (2012). Some consequences of having too little. *Science, 338*, 682–685. https://doi.org/10.1126/science.1222426

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature, 38*(2), 332–382. https://doi.org/10.1257/jel.38.2.332

Thunholm, P. (2004). Decision-making style: Habit, style or both? *Personality and Individual Differences, 36*(4), 931–944. https://doi.org/10.1016/S0191-8869(03)00162-4

Van Swol, L. M., & Sniezek, J. A. (2005). Factors affecting the acceptance of expert advice. *British Journal of Social Psychology, 44*(3), 443–461. https://doi.org/10.1348/014466604X17092

Whitney, P., Rinehart, C. A., & Hinson, J. M. (2008). Framing effects under cognitive load: The role of working memory in risky decisions. *Psychonomic Bulletin & Review, 15*(6), 1179–1184. https://doi.org/10.3758/PBR.15.6.1179

Williamson, S. F., Tishkovskaya, S. V., & Wilson (2025). Hybrid sample size calculations for cluster randomised trials using assurance. *Clinical Trials, 22*(5), 517–526. https://doi.org/10.1177/17407745241312635

Yang, R., Li, S., Qi, Y., Liu, J., He, Q., & Zhao, H. (2025). Unveiling users' algorithm trust: The role of task objectivity, time pressure, and cognitive load. *Computers in Human Behavior Reports, 18*, Article 100667. https://doi.org/10.1016/j.chbr.2025.100667

Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes, 93*(1), 1–13. https://doi.org/10.1016/j.obhdp.2003.08.002

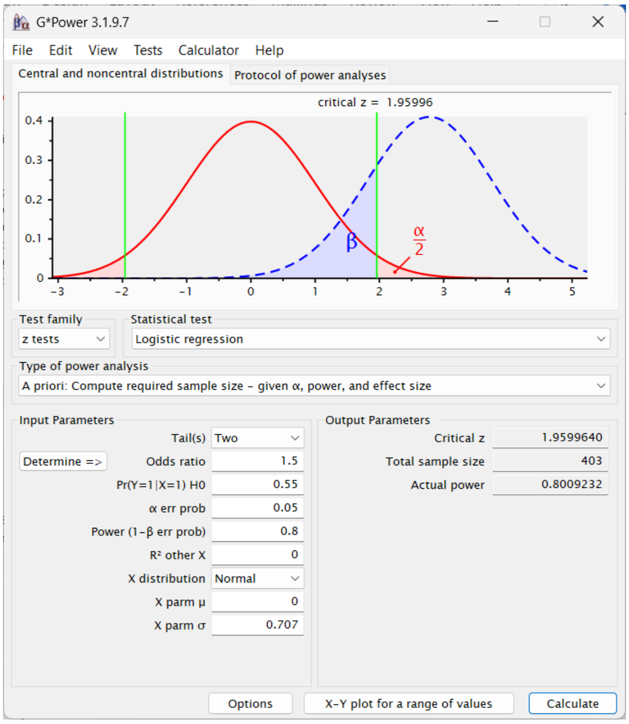
Appendix A

Procedure for the A Priori Power Analysis and Interim Analyses

Firstly, we computed the required sample size as though we would be measuring independent data. Our experiment comprised analyses from multiple researchers; we had to use the input that would yield the largest and safest sample size instead of characteristics of our research. We used G*power 3.1.9.7 (Faul et al., 2009) to perform an a priori power analysis for a two-tailed logistic regression. For determining the effect size, we used one probability and an *OR* instead of two probabilities due to lack of reference effects. The probability p_1 is the probability that under H_0 Y = 1 in the event that X = 1 as described by Faul et al. (2009). For our input X has a normal distribution with parameters $\mu = 0$ and $\sigma = 0.707$. Our significance and power level are $\alpha = .05$ and $1 - \beta = .80$, and after running a sensitivity check, we used $p_1 = .55$ and $OR = 1.5$, which gave us an output of $n_{\text{independent measures}} = 403$, as shown in Figure 1.

Figure 1

*A Priori Power Analysis Output From G*Power 3.1.9.7*



Secondly, interim analyses were performed that adjusted the sample size by accounting for the number of measurements per individual (m_{ind}), using ρ as the intraclass correlation coefficient to compute the design effect (DEFF) that indicates the variation within each cluster (Williamson et al., 2025). We computed DEFF as though each individual is a cluster, as they have multiple measurements, using DEFF = 1 + (m_{ind} – 1)ρ (Rutterford et al., 2015). After running a sensitivity check, we use ρ = .20 and m_{ind} = 16, as we use the analysis of our colleagues without the second block, so DEFF = 4, which indicates the typical variance between the measurements within a participant. We used this ratio to determine the number of effective measurements per individual ($m_{ind/eff}$) by dividing m_{ind} by DEFF to obtain the required $N_{total} = n_{independent\ measures}/m_{ind/eff}$ = 403/4 = 100.75, which we will round up to a multiple of 4 for attrition; hence, we require at least 104 participants for the whole experiment.

Appendix B

Frequency Tables Gender and Age

Table B1

Frequency Table Gender

Gender	*n*	%
Women	64	90.1
Men	7	9.9
Other	0	0
Prefer not to say	0	0

Table B2

Frequency Table Age

Age	*n*	%
18	26	36.6
19	21	29.6
20	6	8.5
21	6	8.5
22	4	5.6
23	3	4.2
24	3	4.2
25	0	0
26	1	1.4
48	1	1.4

Note. Though we recruited participants aged 18–65, only one participant was older than 26. Therefore we left the empty rows for above the age of 26 out to save space. $M = 19.9$, $SD = 3.9$.

Appendix C

Instructions at the Start of the Decision Task – Social Condition

Now you will start the Decision Task.

Your participant code for this part of the experiment is: **XXXXX**
You will complete a simple investment task where you make decisions about money.
Please read the instructions carefully.
Press SPACE to continue.

Task Instructions

In each round, you will be paired with another participant. You will see their participant code and they will see yours (e.g., KM47, TL23). You will interact with a different participant each round.

You have 2 options. You can choose:

- **OUT**: You keep your €10, the other participant gets €0. The round ends.
- **IN**: Your €10 becomes €40 and goes to the other participant. Then the other participant decides:
- **If they choose OUT**: They keep all €40. You get €0.
- **If they choose IN**: They split the €40 with you. You each get €20.

Press SPACE to continue.

Additional Information

- You will make several choices, divided in 2 part (blocks).
- You will be paired with a different participant each round.
- Before your decision, you will receive information about the previous behavior of the partner's in rounds with others. For example, you may see "IN", "IN", "OUT" as the partner's history. This means that on 3 previous trials this was the partner's actual behavior.

- You may also receive a recommendation about what to do. Some recommendations will be generated by an AI agent (ChatGPT) that has been trained on data from previous participants.
- Other recommendations will be provided by other participants in this study. You can use this information as you like. You will also be asked for some recommendations.

Press SPACE to continue to comprehension check.

Comprehension Check

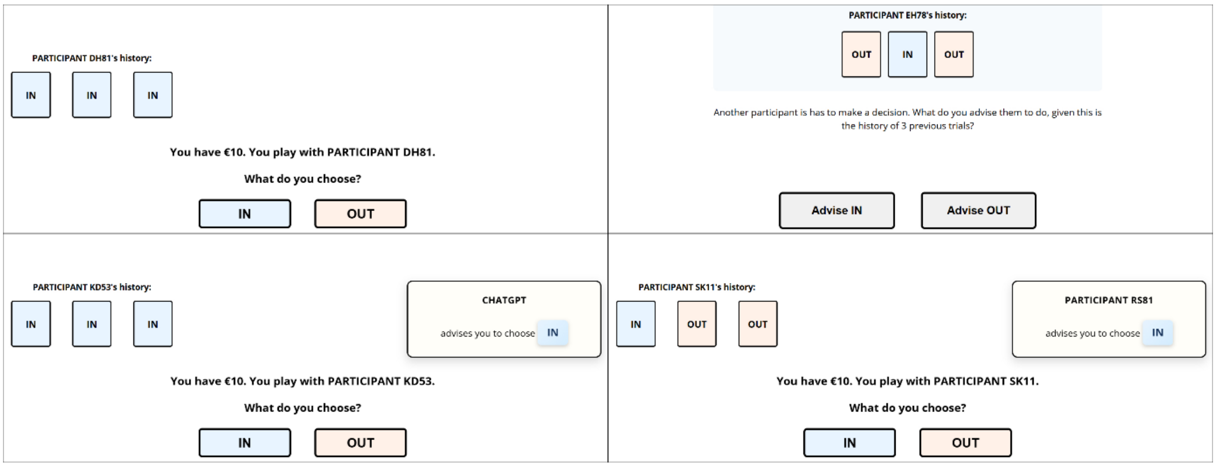
- **Question 1 of 3**: If you choose IN and the other participant chooses OUT, how much money do you get? *Correct answer*: **€0**
- **Question 2 of 3**: If you choose IN and the other participant chooses IN, how much money do you get? *Correct answer:* **€20**
- **Question 3 of 3**: If you choose OUT, how much money do you get? *Correct answer:* **€10**

Appendix D

Visual Design Social Decision Task

Figure D1

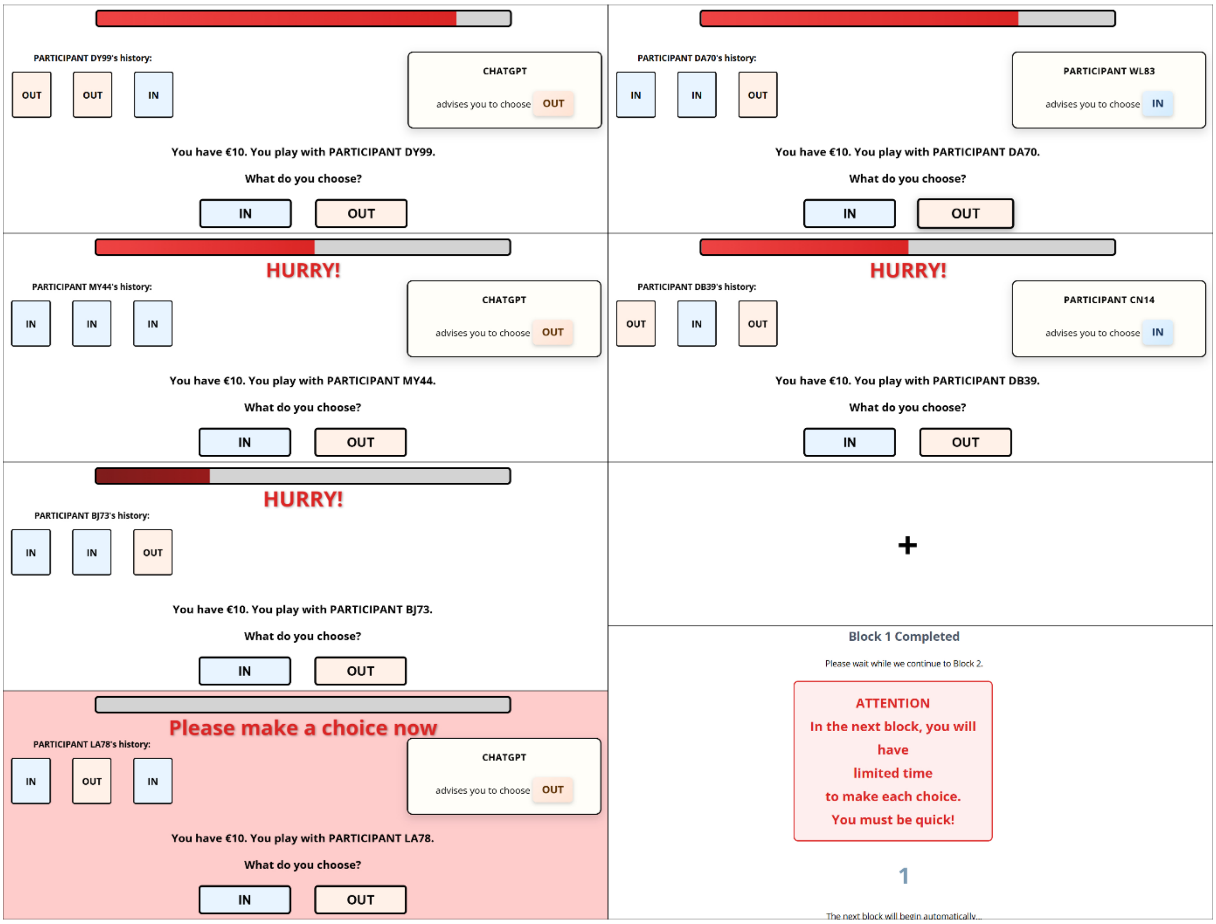
Visual Design First Block



Note. All four trials are from the low-load condition. Up-left: one of the first six *baseline trials* that did not contain advice. Up-right: one of the six trials where participants had to give advice. Down-left: one of the eight trials in the ChatGPT × low-load condition. Down-right: one of the eight trials in the human × low-load condition.

Figure D2

Visual Design Second Block and In-Between Screens



Note: All the displayed trials from the second block are from the high-load condition. From top to bottom: the first image is one of the eight trials in the ChatGPT × high-load condition where a timer running down from 7 s was added; the second image shows the "HURRY" pop-up sign that appears after 3 s; the third image is one of the first six baseline trials without advice; and the fourth image presents the display after the timer has run out.

Right column from top to bottom: the first image is one of the eight trials in the human × high-load condition; the same goes for the second image, where 3 s have passed and the "HURRY" pop-up sign has appeared; the third image is the inter-trial screen that is presented between each trial; and the fourth image is the warning time constraint screen that precedes the second block.

Appendix E

Methodological Skills Framework

Table X

Distribution of Points on the Methodological Skills Framework

Skills class	Level - points	Description
Material design	Medium – 2 points	Selection and/or minor adaption of existing materials to fit the research context (e.g., change stimuli or wording), or coding a simple task.
Data collection	Medium – 2 points	Online data collection from convenience samples (e.g., university students) using standardised survey platforms or simple online experimental paradigms.
Data analysis	Medium – 2 points	Intermediate data analysis, such as mediation, moderation or multiple regression (e.g., including at least two independent variables and one dependent variable)
Open science	Basic – 1 point	Providing a justified sample size rationale (e.g., power analysis). For existing data: sensitivity power analysis. Sharing output and data.

Note. The descriptions are quotes exactly as they are presented in the course manual

Appendix F

Introduction and Information Letter

Welcome to our experiment!

In what follows you will read information about this study, you will be asked for your informed consent, you will engage in a decision task, and you will answer a series of questions.

IMPORTANT: Please make sure to go through until the very last page of this experiment, where you are told it is ok to close the page, in order to receive your SONA participation points!

INFORMATION LETTER

for participation in the scientific research: *Decision Task*

1. General information

1a. Introduction/Objective of the research

This research is conducted by the Behavioural Science Institute, specifically by a group of psychology students as part of their third-year course Research Project 3, with the support of Dr. Catalina Ratala. The aim of the project is to understand how people make choices across different settings. The data collected in this research will be used for academic publication.

1b. The research

In this online study you will do a short task and then you will answer a series of questions, for which there are no right or wrong answers. In the task, you will make a series of simple investment decisions. In each round you will receive an endowment and choose whether to keep it or invest it. Your total participation time will be maximum 30 minutes. There are no risks associated with your participation. The study has been reviewed by the Ethics Committee Social Sciences (ECSS) of the Radboud University (ECSW-2025-109).

2. Privacy

2a. Use of your personal data

For this study we will collect, use, and store personal data. This refers to information through which you can be directly or indirectly identified as an individual. We will collect your age and gender for the demographics of our sample.

2b. Confidentiality of your data and data processing

The information you provide for research purposes is handled carefully and only accessible to authorized personnel. Any personal data collected remains confidential and encrypted. No one else will receive data that can be traced back to you. Please check the annex for more information about right of access by supervisory authorities to inspect the research's compliance with ruling guidelines and additional privacy guidelines applicable.

2c. Data retention period

The consent form you agreed to and your anonymised research data will be retained for 10 years after completion of the research (1st December 2035). After study completion data are stored anonymously and thus cannot be traced back to you.

3. Voluntary participation

Your participation in this research is entirely voluntary. If you decide not to participate in this study, you will have no adverse consequences. Also, during the experiment, you always have the right to withdraw your consent and terminate your participation without any negative consequences for you.

4. Compensation or benefit

Participation in this study will be compensated with study points (SONA points), namely 0,5 ppu.

5. Contact information

We are interested in your experiences as a participant. Should you want to, you can complete an anonymous evaluation online (in Dutch / in English). Should you have any questions, comments, or concerns about this study, please contact us (*e-mail address removed*). Should you prefer to discuss with someone not involved in this study, you can contact the BSI Research Data Officer (*e-mail address removed*), who is an independent point of contact for general research-related matters.

With kind regards,

Catalina Ratala,

e-mail address removed

Behavioural Science Institute,
Radboud University
Thomas Van Aquinostraat 4
6525GD Nijmegen, Netherlands

Annex 1 - Right of access by supervisory authorities to inspect the research's compliance with ruling guidelines

Some individuals and entities must have access to your personal and research data for monitoring purposes. This is necessary to ensure that the research is conducted properly and reliably. The individuals and entities who may access your data for monitoring purposes include: authorized personnel within Behavioural Science Institute or Radboud University (such as a dean, director, or data manager), as well as (inter)national supervisory authorities (such as the Dutch Data Protection Authority and the National Board for Research Integrity). They will keep your data confidential. You will be asked to provide consent for this access. If you do not wish to grant consent, you cannot participate in the research.

Annex 2 - Additional information on your rights regarding the processing of your personal data

The Radboud University is responsible for compliance with the General Data Protection Regulation (GDPR) in the processing of your personal data. The researcher ensures that your privacy and conditions attached to it are safeguarded and adheres to the Dutch Code of Conduct for Research Integrity and university policies for storage and management of personal and research data. You always have the right to withdraw your consent for the processing of your personal data. Your personal data will then be deleted. The Privacy Statement of Radboud University can be found at: https://www.ru.nl/en/about-us/policies-and-regulations/privacy-statements For questions about your privacy, you can contact the Local Privacy Officer of the Faculty of Social Sciences at *e-mail address removed*. For general inquiries, you can contact the Data Protection Officer of Radboud University via *e-mail address removed*. More information about your rights in the processing of your personal data can be found at https://www.ru.nl/privacy/english/protection-personal-data/data-subjects-rights/ and on the website of the Dutch Data Protection Authority (*e-mail address removed*).

Appendix G

Consent Form

Consent form for participation in scientific research for the course Research Project 3: Decision making.

I herewith confirm that:

- I have been informed that the current study is conducted by psychology students as part of their third-year course Research Project 3;
- I have been satisfactorily informed about the study in writing;
- I have read the written information;
- I have been given the opportunity to ask questions about the study;
- my questions have been answered satisfactorily;
- I have been given ample opportunity to think carefully about participating in the study;
- I participate in the study entirely on a voluntary basis.

I understand that:

- I have the right to withdraw my consent to participate at any time during the study without having to state reasons and without fear of adverse consequences;
- my personal data are processed in accordance with the applicable European privacy regulations;
- my personal data are processed in accordance with the privacy statement of Radboud University (https://www.ru.nl/english/vaste-onderdelen/privacy-statement-radboud-university/);
- the tests and questionnaires used are not medical / clinical tests. Since the data are collected anonymously, the researchers cannot inform me about scores that may be of personal clinical interest.

I agree that:

- the online consent form is kept until the end of September 2026 at the latest;
- my personal and research data within this research are obtained for education purposes and will be kept until December 2035 at the latest and will be available for appraisal, verification and audits until that date;
- supervisory authorities may inspect my personal and research data for the purpose of auditing research.

Appendix H

Closure and Written Debriefing

In closing...

You have reached the end of this experiment.

IMPORTANT: make sure to go through until the final page to get the SONA points.

With this study we aim to understand how individuals incorporate advice into their daily choices. Specifically, we want to see if receiving advice from another person compared receiving advice from an AI agent impacts ones' risk taking in social and non-social contexts. For this reason, the advice you saw on the screen was all carefully generate by the experimenters. This means that you did not receive real time advice on your choices for other people. We apologize for misleading you, but this was an important aspect of our research. Should you want more information about the study or should you desire to withdraw your participation, please do not hesitate to contact Catalina Ratala at *e-mail address removed*.

Additionally, we were interested in how self-efficacy, the belief in one's ability to handle challenges and reach goals, may influence the extent to which people follow advice from these different sources. Besides this, some participants completed the task under time pressure to study the effect of cognitive load on advice taking, while others performed the task without such pressure. We also measured neuroticism, a personality trait related to emotional stability, to explore how individual differences in personality may be linked to decision-making and reliance on advice.

By comparing these conditions, we can learn more about how individuals perceive, and trust AI systems compared to human advisors, especially when facing risky or uncertain choices. These insights can help us understand how advice functions in human–AI interactions and how to improve the design of AI systems that support decision-making.

We would like to ask you not to share the information about this study with others, to ensure that new participants are not familiar with the purpose of our study.

At the start of the study we could not explain the information above completely, because we expected that knowing in advance about the study could influence the results. Should you wish to withdraw your consent for the study participation, please contact the researchers at *e-mail address removed* or *e-mail address removed*.

We want to reassure you that the information we collected will be processed anonymously. This means that the results cannot be traced back to you. The consequence of this is that we cannot inform you about your personal results after the study has been completed. In addition, this means that we cannot inform you about scores that might be of personal clinical significance to you.

Did this study unintentionally prompted unpleasant feelings, thoughts or insecurities for you, or do you have any concerns about your well-being after completing this study? Then, please contact the study advisor, the student psychologist or your general practitioner for an appointment.

Do you want to know more about the resources Radboud University offers to support students with their well-being? If so, please click on the following link: https://www.ru.nl/currentstudents/during-your-studies/overview-counsellors-advisors/student-support/

If you wish to be informed about the results of this study, then please let us know via *e-mail address removed*. Also, for questions, remarks or concerns about this study, you can contact us via *e-mail address removed*. If you have questions, remarks or concerns that you rather not share with the researchers in charge, you can contact the coordinator of Research Project 3 (Willemien de Kleijn, *e-mail address removed*).

Thank you very much for participating in our study!
Kind regards,

Anastasia Peijs, Maud Inghels, Manisha Ramadhin, Joep Voorter, Jip Damen & Julidé Wagner
Students Research Project 3
Supervisors: Catalina Ratala & Alan Sanfey
Radboud University

Appendix I

Descriptive Statistics of Key Variables

Table I1

Advice-Taking Rates per Condition

Condition	m_i	*M*	*SE*	95% CI	
				LL	*UL*
Low-load × ChatGPT	8	.412	.415	.024	.953
Low-load × human	8	.419	.417	.024	.954
High load × ChatGPT	8	.406	.413	.023	.952
High load × human	8	.424	.418	.025	.955
Low-load, both SoA	16	.415	.415	.024	.953
High-load, both SoA	16	.415	.415	.024	.953
ChatGPT, both CL	16	.409	.414	.024	.952
Human, both CL	16	.421	.417	.025	.954
Both CL, both SoA	32	.415	.415	.024	.953

Note. $N = 71$. m_i = number of measurements per individual; CI = confidence interval; *LL* = lower limit; *UL* = upper limit; SoA = source of advice; CL = cognitive load; $m_i = m_{total}/N$. Advice-taking rates are displayed as the means of all relevant measurements, where conforming to the advice was registered as 1 and rejecting the advice as 0.

Appendix J

Frequency Tables Advice Taking

Table J1

Frequency Table of Advice Taking in Every Condition

Number of taken advices	Low-load				High-load			
	Human		ChatGPT		Human		ChatGPT	
	n	%	*n*	%	*n*	%	*n*	%
2/8	2	2.8	1	1.4	0	0	1	1.4
3/8	3	4.2	5	7.0	9	12.7	4	5.6
4/8	28	39.4	30	42.3	28	39.4	30	42.3
5/8	26	36.6	22	31.0	22	31.0	20	28.2
6/8	10	14.1	5	7.0	7	9.9	11	15.5
7/8	1	1.4	7	9.9	4	5.6	3	4.2
8/8	1	1.4	1	1.4	1	1.4	2	2.8
	Frequency rates of advice taking from other conditions							
	Both levels of CL				Both levels of SoA			
	Human		ChatGPT		Low-load		High-load	
	n	%	*n*	%	*n*	%	*n*	%
5/16	1	1.4	1	1.4	1	1.4	0	0
6/16	3	4.2	0	0	0	0	2	2.8
7/16	2	2.8	4	5.6	5	7.0	7	9.9
8/16	20	28.2	21	29.6	20	28.2	15	21.1
9/16	17	23.9	18	25.4	18	25.4	20	28.2
10/16	14	19.7	11	15.5	12	16.9	11	15.5
11/16	7	9.9	6	8.5	6	8.5	9	12.7
12/16	3	4.2	5	7.0	5	7.0	2	2.8
13/16	2	2.8	1	1.4	2	2.8	1	1.4
14/16	2	2.8	2	2.8	0	0	3	4.2
15/16	0	0	1	1.4	2	2.8	1	1.4
16/16	0	0	1	1.4	0	0	0	0

Note. $N = 71$. CL = cognitive load; SoA = source of advice. For the latter two conditions, all participants took the advice at least five times, so the empty rows are excluded to save space.

Descriptive statistics for the six conditions that are displayed in the table: Low-load × human (*M* = 4.6, *SD* = 1.0); low-load × ChatGPT (*M* = 4.7, *SD* = 1.2); high-load × human (*M* = 4.6, *SD* = 1.1); high-load × ChatGPT (*M* = 4.7, *SD* = 1.1); both CL × human (*M* = 9.3, *SD* = 1.8); both CL × ChatGPT (*M* = 9.5, *SD* = 2.0); low-load × both SoA (*M* = 9.4, *SD* = 1.8); high-load × both SoA (*M* = 9.4, *SD* = 1.9).

Table J2

Frequency Table Advice Taking Across all Trials

Number of taken advices	*n*	%
13/32	1	1.4
14/32	2	2.8
15/32	2	2.8
16/32	15	21.1
17/32	10	14.1
18/32	12	16.9
19/32	9	12.7
20/32	3	4.2
21/32	3	4.2
22/32	4	5.6
23/32	4	5.6
24/32	2	2.8
25/32	0	0
26/32	2	2.8
27/32	1	1.4
28/32	0	0
29/32	0	0
30/32	1	1.4
31/32	0	0
32/32	0	0

Note. *N* = 71, *M* = 18.7, *SD* = 3.3. From all 32 trials, all participants took the given advice at least 13 times, so the empty rows are excluded to save space.