

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Gao, Y;Garagnani, M

Title:

Improving risky choices: The effect of cognitive offloading on risky decisions

Date:

2025-04-01

Citation:

Gao, Y. & Garagnani, M. (2025). Improving risky choices: The effect of cognitive offloading on risky decisions. *Journal of Risk and Uncertainty*, 70 (2), pp.105-128. <https://doi.org/10.1007/s11166-025-09451-z>.

Persistent Link:

<https://hdl.handle.net/11343/362196>

License:

[CC-BY](#)



# Improving risky choices: The effect of cognitive offloading on risky decisions

Yihong Gao<sup>1</sup> · Michele Garagnani<sup>1</sup> 

Accepted: 11 February 2025 / Published online: 12 April 2025  
© The Author(s) 2025

## Abstract

Economic choices are difficult and computationally complex. To make good decisions one needs to integrate different variables and resolve several trade-offs. We study whether providing external tools which help reducing the computational complexity of choices improves risky decisions. We find mixed evidence for the effectiveness of these tools, as people tend to rarely use them. Overall, cognitive offloading tools make people less risk averse and better at updating beliefs, but they do not decrease irrational behavior, in the sense of the proportion of dominated choices, violations of expected utility, and sub-optimal budget allocations. These results suggest that simply providing external tools, which make choices (computationally) easier, but where using these aids is discretionary, has a limited impact on improving people's choices under risk.

**Keywords** Risk · Rationality · Offloading · Decision quality

**JEL Classification** D01 · D81 · D91

## 1 Introduction

People do not generally decide optimally. There is a veritable catalog of experimental and empirical works showing people do not behave as normative models prescribe (e.g., Tversky, 1969; Kelman et al., 1998; Hey et al., 2009; Taylor, 2013; Dumm et al., 2020; Gillenkirch & Velthuis, 2023). There are several explanations for why this happens, and people might be optimally adapting choice quality based on their cognitive restraints and expected reward (e.g., Glimcher, 2022). However, in this work we focus on the fact that these shortcomings have been mostly studied in controlled experimental settings where participants are traditionally prevented from using external tools, e.g., calculators. However, most decisions, especially important ones, are made when nothing prevents people from using external helpful aids. It is therefore not clear whether

---

✉ Michele Garagnani  
michele.garagnani@unimelb.edu.au

<sup>1</sup> Centre for Brain, Mind, and Markets, Department of Finance, University of Melbourne, Melbourne, Australia

people might actually still display the biases and shortcomings reported in the experimental literature when they can rely on helpful tools, or whether behavior is somehow better outside laboratory settings thanks to the use of these tools. In this sense, this work aims to study whether people's decisions among risky prospects improve in quality and adhere more to normative prescriptions when cognitive offloading tools are readily available.

Cognitive offloading refers to the use of mediums to overcome limitations in a decision maker's cognitive capacities (Risko & Gilbert, 2016). Everyday examples of cognitive offloading tools include the use of a calculator to assist in arithmetic calculations or the reliance on digital reminders to store to-be-remembered information. Cognitive offloading has been at the center of several recent publications, but it is mostly focused on memory tools (e.g., Boldt & Gilbert, 2019; Storm & Stone, 2015; Sparrow et al., 2011; Morrison & Richmond, 2020; Burnett & Richmond, 2023), particularly on short-term memory (Risko & Dunn, 2015), mental rotation (Chu & Kita, 2011), and learning (Costa et al., 2011). However, we are mostly interested in offloading of computations, as this might be more likely to impede good choices in the economically-relevant settings where biases have been reported, i.e., lottery choices.

Studies that investigate the effect of cognitive offloading in conjunction with economically-relevant decisions are however rare. The closest related works are those which study the role of experts advice, for example on financial decision (e.g., Engelmann et al., 2009). In this sense delegating decision making to experts is essentially offloading decisions to other agents. Hence this type of study focuses more on what drives people to delegate their choices than whether these (offloading) tools directly improve people's choices. To the best of our knowledge, there is currently no work on how cognitive offloading shapes the quality of decision making under risk.

Choices under risk are of preferential nature, and hence objective metrics for decision quality are harder to define. However, we focus on risky choices because this is one of the most common area of research in decision making, and it is where several biases and shortcomings have been traditionally documented (e.g., Allais, 1953; Kahneman & Tversky, 1979). In particular, we focus on some of the most prominent examples of irrational choices as well as some previously proposed metrics of choice quality within this domain.

Firstly, we study whether risk attitudes are affected by the use or availability of cognitive offloading tools. We do so, by using the classical paradigm of lottery choices, but also a more ecologically-valid card-betting task (Alós-Ferrer & Garagnani, 2022a, b). While whether individuals are more or less risk averse is not a direct indicator of choices quality, it is nevertheless interesting to study whether people's choices under risk are affected by the use of calculators, computers, or other devices which make computations easier.

We also study choice inconsistency. While people should choose consistently what they prefer, it is often the case that even when exposed to the same choices multiple times, people tend to choose differently (e.g., Hey & Orme, 1994; Ballinger & Wilcox, 1997; Agranov & Ortoleva, 2017; Alós-Ferrer & Garagnani, 2021b). Many works have reported anomalies where human decision making violate some form of consistency (e.g., Tversky, 1969; Alós-Ferrer et al., 2024b). However, the relationship between cognitive offloading and the inconsistency in risky decisions is still unex-

plored. In particular, whether cognitive offloading tools, especially those which help with computations, e.g., calculators, improve the consistency of risky decisions is an empirical question. Further, an interesting question is whether this eventual improvement is a discrete jump or is mediated by other variables, e.g., the subjective value of the options. Alós-Ferrer and Garagnani (2022a,b) find that, for risky decisions, inconsistent choices are less frequent when the (subjective) value of the options are far apart from each other. Some previous works suggest that people might be aware of whether they are close to indifference (e.g., Cettolin & Riedl, 2019; Halevy et al., 2023). In these situations people might be more inclined to use offloading tools to help their decisions.

Another objective measure of decision quality even under risk is the proportion of first-order stochastic dominance (FOSD) choices. A prospect is first-order stochastic dominated when there is another prospect that consistently yields higher outcomes across all values, i.e., delivers at least the same monetary outcomes with higher probability or more money with at least the same probability. FOSD is interesting for our case because it allows to compare the quality of choices independently of the decision maker preferences, i.e., there are no preferences which would justify choosing dominated options, other than, e.g., trembling-hand mistakes.

We also study behavior in budget-choice allocation tasks. Among the topic of risky choices, budget-choice tasks are of particular interest mainly because of the wide range of choices available for selection following multiple budget constraints. For example, these methods of choice are often used to test the consistency of subjects' choices with GARP (e.g., Choi et al., 2014). In a budget-choice task, the participants are typically asked to allocate their wealth between a riskier asset and a less risky asset. A common setup for budget-choice task is by using sliders given the relatively intuitive and user-friendly interface for elicitation (e.g., Gneezy & Potters, 1997; Choi et al., 2007; Hey & Pace, 2014; Halevy et al., 2018; Kurtz-David et al., 2019). Garagnani (2023) showed there is a high degree of sub-optimal behaviors in budget-choice tasks by designing the tasks as having a global optimal solution for all (risk averse) participants regardless of their levels of risk aversion. The high degree of sub-optimality motivates us to investigate if cognitive offloading tools potentially lead to better choices in such tasks.

Among the various forms of irrational behavior under risk, we focus on the certainty effect (e.g., Allais, 1953; Kahneman & Tversky, 1979). This is one of the most prominent violations of Expected Utility Theory (EUT) and is the empirical observation that, when choosing between a sure but small amount and a larger but risky amount, most participants choose the safe alternative, but when all probabilities are scaled down by a common ratio, most people tend to choose the riskier option (Kahneman & Tversky, 1979). This type of behavior is incompatible with any form of expected utility maximization and hence it is classified as a form of irrational behavior.

In an online experiment where we randomly assign participants to either have no tools or can use cognitive offloading tools, we find a limited effect of these external help in risky choices. This limited effect seems to be driven by the fact that people tend to scarcely rely on the provided tools. When people use them, they are less risk averse and better at updating their beliefs. However, providing people with offloading tools does not make them more rational, in the sense of displaying less of the certainty effect,

choose less dominated options, or choose more optimal allocations in budget-choice tasks.

The rest of the paper is structured as follows: We introduce the design of the experiment and then report our pre-registered hypotheses. We then report the results of the study separately for each task: Binary-lottery choices (subdivided in the analysis of risk attitudes, dominated choices, and support for the certainty effect), budget-choice task, and a belief updating task. Finally, we explore what drives the reliance on the offloading tools.

## 2 Experimental design and methods

We conducted an online study on Prolific with a program coded in Unity. At the beginning of the study participants were informed about the general nature of the study and we collected their informed consent to participate. They were also informed that their data would be treated completely anonymously, and that they could leave the study at any moment in case they wanted to withdraw their consent. They then received more detailed instructions informing them about the structure of the study, the payment procedures, and the exchange rate between the Experimental Currency Units in which outcomes were represented and Pounds (10 ECUs = 1 Pound).

Figure 1 depicts the structure of the experiment. Participants were randomly assigned to one of two between-subjects conditions, OFFLOADING vs. NO-OFFLOADING, which corresponded to receiving (cognitive offloading) help or not, respectively. The form of the offloading mechanism depended on the type of task. The study was composed of three main parts, lottery choices, budget-choice allocation task, and a card-betting task, plus a questionnaire. In particular, we asked them to perform the Numerical Literacy Scale task (Lipkus et al., 2001), which allow us to have a proxy for their mathematical skills.

Choices were incentivized using the random incentive mechanism (Cubitt et al., 1998). To determine a participant's payoff, one choice for each part of the study was randomly selected. The option chosen by the participant was played out, and the participant received the realized outcome as payoff. In addition, participants received a show-up fee of 3.5 Pounds for an average total remuneration of 5 Pounds, which is in line with Prolific standards. Average time of completion is 21.7 minutes for the OFFLOADING group and 20.1 minutes for the NO-OFFLOADING group.

### 2.1 Lottery choices

The first part of the study required participants to make binary-lottery choices. In the lottery choices the offloading tool took the form of a calculator which was provided to participants on each decision screen (see Fig. 3 for screenshots). The 26 lottery pairs are of three types, each has a different function and hence we analyze them separately. See Table 4 for the list of lottery pairs.

A subset of six lottery pairs reproduced the certainty effect, in particular the common-ratio effect (CR) (Kahneman & Tversky, 1979). The construction of these

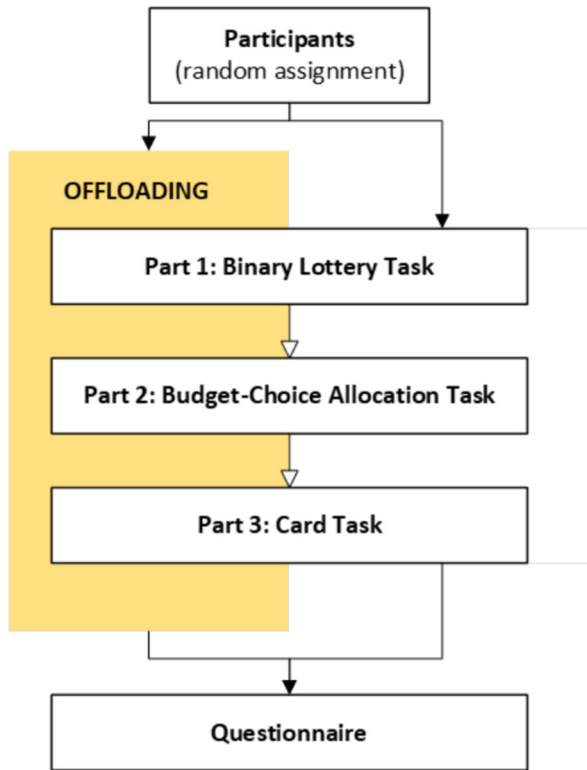


**Fig. 1** Structure of the experiment

pairs is taken from Alós-Ferrer et al. (2024a) and is as follows. Given monetary outcomes  $X, Y$  with  $X > Y > 0$ , a probability  $p$  with  $0 < p < 1$ , and an integer  $k \geq 1$ , define the pair of lotteries  $S(k)$  and  $R(k)$  by

$$S(k) = \begin{cases} Y & \text{with prob. } 1/k \\ 0 & \text{otherwise,} \end{cases} \quad R(k) = \begin{cases} X & \text{with prob. } p/k \\ 0 & \text{otherwise.} \end{cases}$$

Note that  $R(k)$  is always riskier than  $S(k)$ , hence the latter is “safer.” The CR effect involves a preference reversal between two related pairs of lotteries. The *certainty pair* is given by setting  $k = 1$  above. Then  $S_C = S(1)$  is a sure gamble, while  $R_C = R(1)$  has probability  $p$  of a non-zero outcome. The *normal pair* is derived from the certainty pair by scaling down the probabilities of non-zero outcomes by the ratio  $1/k$ . That is, given some  $k > 1$ , one defines  $S_N = S(k)$  and  $R_N = R(k)$ , none of which involve a certain outcome. For instance, letting  $X = 21$ ,  $Y = 16$ , and  $p = 0.8$ , we obtain the certainty pair with  $S_C = (1, 16)$  and  $R_C = (0.80, 21)$ . Dividing the winning probabilities by  $k = 4$ , we obtain the normal pair with  $S_N = (0.25, 16)$  and  $R_N = (0.20, 21)$ . Under EUT, any decision maker who chooses  $S_C$  in the certainty pair must also choose  $S_N$  in the normal pair (and analogously for  $R_C, R_N$ ). Empirically

(Ruggeri et al., 2020; McGranaghan et al., 2024), a large number of people choose  $R_N$  and  $S_C$ , hence contradicting EUT. Therefore, we define the proportion of choices of the type  $(S_N, R_C)$  and  $(R_N, S_C)$  over all choice in the double pairs as the proportion of certainty-effect behavior.

Four of the lottery pairs involved first-order stochastic dominance. That is, one of the lottery gave either less money with the same probability or the same amount of money with less probability. In this sense, everybody, independently of their risk attitudes should choose the non-dominated lottery.

The rest of the pairs were designed according to optimal-design theory, see Moffatt (2015) and Garagnani (2023) for the construction of the lottery pairs and their rationale. This set of lotteries allows us to estimate subjects' risk preferences precisely. In particular, individual risk preferences are estimated for each participant from their binary choices between the lotteries using a maximum likelihood procedure. The procedure followed the approach described in Moatt (2015, Chapter 13). In particular, we implement the same estimation detailed in Alós-Ferrer and Garagnani (2022a) (see also Alós-Ferrer et al., 2024b; Alós-Ferrer & Garagnani, 2021b; Garagnani, 2023) and redirect to those papers for a detailed description of the estimation procedure. We estimated an additive random utility model (e.g., Thurstone, 1927; Luce, 1959; McFadden, 2001) with CRRA utility function. We estimated individuals' risk attitudes from the binary-lottery task and use them to predict choices and classify participants in the rest of the experiment (e.g., for the card-betting task and the budget-choice task).

## 2.2 Budget-choice allocations

The second part of the study involved budget-choice allocation tasks. We adopted the “slider approach” in which each participant has to complete three sequential budget-choice tasks by moving the corresponding sliders to the positions they desire (see Fig. 4 for screenshots). The positions of the sliders indicate the probabilities and/or outcomes of the underlying assets, which are visible to the participant at all times. Hence, participants choose their preferred risky allocation by trading off the magnitudes of the monetary amounts and their related probabilities. All three budget-choice tasks have a fixed expected value of four ECUs and the payoffs of the assets are jointly determined to preserve the expected value. Each slider was constructed such that the solution to the maximization problem is always a corner solution for any risk averse participant (Garagnani, 2023). That is, each participant who is risk averse should normatively choose a corner solution in all sliders. We use participants' choices in the lottery task to classify them as risk averse or not. The initial budget-choice task is the same for every participant, whereas subsequent tasks are participant-specific, meaning that they were constructed using the outputs from the first slider. In particular, the first task asks people to solve the following maximization problem:

$$\max_{p \in [0.05, 0.95]} p \cdot u\left(\frac{K}{p}\right), \quad (1)$$

by changing the probability  $p$ , subjected to the expected value  $K$ .

The second task asks people to maximize:

$$\max_{q \in [0.01, q^*]} p^* \cdot u(K + (1 - p^*)q) + (1 - p^*) \cdot u(K - p^*q), \quad (2)$$

by changing the outcome  $q$ , subjected to  $K$  and the chosen probability from the first task  $p^*$ .  $q^*$  is the chosen outcome from the first task.

The third and final task asks participants to maximize:

$$\max_{p \in [0.05, p^*]} p \cdot u(K + (1 - p)q^*) + (1 - p) \cdot u(K - pq^*), \quad (3)$$

by changing  $p$ , subjected to  $K$  and the chosen outcome from the first task  $q^*$ .  $p^*$  is the chosen probability from the first task.

All probabilities are restricted to between 0.05 and 0.95 as to avoid the certainty effect, and outcomes are restricted to positive numbers to avoid losses. This design allow us to test again for decision quality, this time among budget-choice tasks. The offloading tool for the budget-choice tasks took the form of a calculator which was provided to participants on each decision screen.

### 2.3 Belief updating in risky choices

The third part of the study asked participants to make choices in a card-betting game which is a different way of representing risky prospects, but also requires updating of beliefs in order to evaluate the options presented. The task is taken from Alós-Ferrer and Garagnani (2022a). There were two decks of cards at the start of each round, each deck contained cards with face values 1–10 either in all clubs (black) or diamonds (red). Participants were then shown two black cards and one red card at random. The participants' task is to decide whether to bet on the third black card drawn would be greater than the red card or to pass. For example, the initial cards drawn can be a black 4, a black 8 and a red 9. The participant might decide to take the bet if they believe the third black card would be greater than the red 9. That is, participants are effectively choosing between lotteries, but these are framed as a card-betting game where to calculate the value of the bets participants needed to update their beliefs according to the provided information.

Betting is costly, placing a bet costs 0.10 Pounds (fixed for all trials), independently of the outcome of the trial. If the participant bets and if the newly extracted black card is strictly larger than the displayed red card, the participant receives the displayed monetary prize (minus the cost). Otherwise, the payment is zero (resulting in a net loss equal to the cost of betting). If the participant does not bet, there is neither a payment nor a cost (but the new black card is still extracted and the participant observes the outcome). Before a new trial starts, all cards are returned to their respective decks and those are reshuffled. Hence, each trial reflects an independent decision situation.

The set of initial stimuli (red card, first two black cards, and prize) was predetermined and pseudorandomized across trials to achieve adequate stimuli variance with a manageable number of trials. All participants were exposed to the same set of stimuli,

albeit in randomized order. Red cards were extracted in such a way that there was always some probability of winning, so as to avoid trivial decisions. Hence, there were eight possible distinct probabilities of winning, ranging from 12.5% to a sure win. Prizes ranged from 8 to 112 pennies and were determined as deviations from the fair prize in 10% increments, ranging from 50% less than the fair prize to 50% more. Each participants completed 44 trials of this task. Only one randomly selected trial was paid if this task was selected to be relevant to determine participants' earnings.

The offloading tool for this part of the study was different from the other tasks. Under offloading, participants were presented with a screen showing the card deck information for the black deck. On this offloading screen, participants can visually identify both the number of winning and losing cards and their respective face values, i.e., the likelihood of winning the bet. Moreover, the two types of cards (winning/losing) were contrasted with highlights using complementary-color (see Fig. 5 for screenshots). In this sense, this is the only task among the ones we implemented where the use of the offloading task was not discretionary but was always present.

### 3 Hypotheses

We preregistered directional hypotheses for each part task in the experiment, see AsPredicted (187585). The intuition behind the hypotheses is that we expected the cognitive offloading tool to improve decision quality, in the sense of filling the gap between behaviors under normal circumstances (when no offloading is provided) and what people should normatively do. This might also translate into more risk neutral choices as providing participants with a way of offloading their computations might prime them to indeed compute the expected value of the options.

- In the lottery task, we expected participants in the OFFLOADING compared to the NO-OFFLOADING condition to be less risk averse, to display a lower proportion of dominated choices, and to display less common-ratio type behavior.
- In the budget-choice task, non-risk loving participants in the OFFLOADING condition are expected to make more optimal choices compared to those in the NO-OFFLOADING condition.
- In the card-betting task, we expected participants to make less errors in the OFFLOADING compared to the NO-OFFLOADING condition. We define an error as a choices in favor of options with negative expected utility, e.g., deciding to bet when the expected utility of betting is strictly smaller than the expected utility of passing and vice versa.

### 4 Results

We obtained ethical approval by the Research Ethics Unit of the University of Melbourne (Project ID 29953). The study was pre-registered at AsPredicted (187585). A total of 293 subjects (151 females); median age 33, (ranging from 18 to 87) participated in an online study on Prolifics, of which 144 are from NO-OFFLOADING and 149

are from OFFLOADING. Participants were recruited among the general population registered on the platform. The sample size was determined by a power calculation in order to guarantee an a priori power of 0.8, for a non-parametric, between-subject Mann-Whitney test, for a small to medium effect size of 0.3.

#### 4.1 Lottery choices

In the lottery task participants used the cognitive offloading tool (i.e., the calculator) 279 times out of the 3744 trials. Specifically, 40 out of the 144 (27.78%) participants in the OFFLOADING condition used the calculator at least once in this task. As a preliminary analysis, we first explore the relationship between lottery choices, offloading, and expected value maximization. Subjects in the OFFLOADING group make more choices aligned with EV maximization (60.37%) compared to the NO-OFFLOADING group (53.86%) according to a between-subject, non-parametric Wilcoxon-Mann-Whitney test (WMW test,  $N = 293$ ,  $z = 2.053$ ,  $p = 0.040$ ). For this and the other tasks we report the results of a series of robustness analyses which consider only the restricted sample of subjects who actually used the cognitive offloading tool (see Appendix C). We remark that the analysis considering only the restricted sample is likely affected by selection effects. Hence, the results of the restricted sample analyses should be interpreted with further caution. The results of the full and restricted sample are qualitatively the same, with the latter showing stronger effects.

To further investigate the factors contributing to this effect, we perform a series of random-effects panel probit regressions on the probability of a “correct” choice (i.e., a choice in agreement with expected value or expected utility maximization). Model 1 of Table 1 reproduces the results of the non-parametric tests. The treatment effect (dummy variable *Offload*) is positive and significant. This is true even when controlling for the expected value of the options (*EV\_Dist*) and for individual differences, i.e., Age, Gender, and Numerical Literacy (Model 2). Moreover, participants displayed a higher probability of a choice in alignment with expected value maximization (Model 2, Offloading dummy  $\beta = 0.243$ ,  $SE = 0.143$ ,  $p = 0.089$ ). Model 1 of Table 5 shows the effect is stronger for the sub-sample of participants who actually used the cognitive offloading tool. These preliminary results agree with the idea that cognitive offloading improves choice quality among risky decision, in particular that it might nudge people to choose more in line with expected value maximization. This motivates us to investigate further these effects by incorporating individuals’ heterogeneous risk attitudes.

##### 4.1.1 Risk preferences

We observe a lower degree of risk-aversion in OFFLOADING ( $\mu = -2.595$ ,  $SD = 0.173$ ) compared to NO-OFFLOADING ( $\mu = 0.262$ ,  $SD = 0.049$ ; WMW test,  $N = 293$   $z = -11.968$ ,  $p < 0.001$ ).<sup>1</sup> That is, we find support for the idea that subjects

<sup>1</sup> We report the results of the estimation of a Random Utility Model with CRRA utility function of the form  $u(x) = \frac{x^r}{r}$ . Hence, larger  $r$  indicate lower risk aversion for  $r \in [0, 1]$ .

**Table 1** Random-effect probit regression of the probability of a “correct” choice defined as choosing in agreement with expected value (Models 1 and 2) or utility maximization (Model 3)

| Correct Choice  | Model 1            | Model 2              | Model 3              |
|-----------------|--------------------|----------------------|----------------------|
| Offloading      | 0.214**<br>(0.109) | 0.243*<br>(0.143)    | 0.226***<br>(0.078)  |
| Age             |                    | −0.001<br>(0.006)    | 0.006*<br>(0.008)    |
| Female          |                    | −0.413***<br>(0.143) | 0.002<br>(0.078)     |
| Num.Lit         |                    | 0.122<br>(0.075)     | −0.050<br>(0.041)    |
| EV_Dist         |                    | 2.574***<br>(0.094)  |                      |
| EU_Dist         |                    |                      | 68.809***<br>(8.152) |
| Constant        | −0.020<br>(0.076)  | −0.899***<br>(0.306) | 0.036<br>(0.165)     |
| Log L           | −2777.51           | −2139.68             | −2552.55             |
| Wald test       | 3.83*              | 759.55***            | 94.38***             |
| Obs. (trials)   | 4688               | 4464                 | 4464                 |
| Obs. (subjects) | 293                | 279                  | 279                  |

Note: Models 3 and 6 include only those participants who used the cognitive offloading tool at least once. Robust standard errors in brackets  
 \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

are more risk neutral (less risk averse) when provided with cognitive offloading tools, especially when they make use of them.

In agreement with the previous result on expected value maximization, we also find that when defining a “correct” choice as a decision in agreement with expected utility maximization, OFFLOADING (68.84%) compared to NO-OFFLOADING (59.61%) leads to a higher proportion of correct choices (WMW test,  $N = 293$ ,  $z = 3.075$ ,  $p = 0.002$ ). We reproduce these results using regressions analyses. Model 3 in Table 1 shows that the probability of a correct choice, defined as following expected utility maximization, is higher under OFFLOADING even when we control for Age, Gender, and Numerical Literacy and for the subjective choice difficulty ( $EU\_Dist$ , Offloading dummy  $\beta = 0.226$ ,  $SE = 0.078$ ,  $p < 0.001$ ).

#### 4.1.2 Dominated choices

We find no support for a significant difference in the proportion of dominated choices between OFFLOADING (3.47%) and NO-OFFLOADING (2.35%, WMW test,  $N = 293$ ,  $z = 0.361$ ,  $p = 0.718$ ). The absence of a treatment difference goes against our hypothesis, but it is not a surprising result. Dominated choices are easy to spot and error rates related to violation of first-order stochastic dominance are rare. Hence, there is little space for an effect of our cognitive offloading manipulation. Generally, the result regarding dominated choices suggests that this type of irrational behavior

cannot be reduced using only a computational-type offloading tool, but it might be due to noise in perception (e.g., Garagnani & Vieider, 2002) or trembling-hand mistakes (e.g., Loomes et al., 2002).

#### 4.1.3 Certainty effect

We investigate the proportion of behavior compatible with the common ratio effect between treatments. The OFFLOADING group did not display a significantly different proportion of certainty-effect behavior (49.54%) compared to the NO-OFFLOADING group (47.65%, WMW test,  $N = 293$ ,  $z = 0.465$ ,  $p = 0.642$ ). This result goes against our hypothesis of a treatment difference, but it is interesting, as it hints at the idea that expected utility violations of this kind might not be due to complexity in calculations, but might be a type of systematic preference (e.g., Alós-Ferrer et al., 2024a).

### 4.2 Budget-allocation task

In the budget-choice task the main dependent variable of interest is the proportion of optimal choices (i.e., corner-solution choices among risk averse participants). The risk-preference estimation based on the lottery choices classifies 28 (9.56%) of the participants as risk seeking, we exclude them from this part of the analysis. However, the results are qualitatively unchanged when we include all subjects in the analysis. Within the budget-choice task, a total of three participants (2.40%) in the OFFLOADING treatment used the calculator at least once.<sup>2</sup>

Figure 2 shows the distributions of answers in the three budget-allocation tasks. We classify choices within the 5% range of the optimal corner solutions on the slider as optimal choices, in order to account for any visual discrepancies or imprecisions in the responses. Comparing the two treatment groups, we find no statistically significant differences between OFFLOADING and NO-OFFLOADING for all three sliders. In particular, the individual proportion of optimal choices in the first slider for OFFLOADING (26.40%) is not significantly different from the one in NO-OFFLOADING (22.86%, WMW test,  $N = 265$ ,  $z = 0.668$ ,  $p = 0.504$ ). Similarly for slider two (OFFLOADING: 36.80%, NO-OFFLOADING: 38.57%; WMW test,  $N = 265$ ,  $z = -0.296$ ,  $p = 0.767$ ) and slider 3 (OFFLOADING: 28.00%, NO-OFFLOADING: 27.14%, WMW test,  $N = 265$ ,  $z = 0.156$ ,  $p = 0.876$ ).

The lack of significant differences in behavior between OFFLOADING and NO-OFFLOADING goes against our hypothesis and might suggest that people do not understand how to use budget-choice allocation task, as previously indicated in the literature (e.g., Garagnani, 2023). However, by providing participants with a cognitive offloading tool, we further provide evidence that this lack of understanding is probably not due to the computational complexity of the task, as this is reduced by the provided tool. However, the general issue is that people almost never use the provided tool in this task, hence the eventual effect of the cognitive offloading manipulation is mechanically limited.

<sup>2</sup> twelve participants in OFFLOADING used the offloading tool at least once (4.10%), but nine of them are classified as risk seeking and hence are excluded from the analysis reported in this subsection.

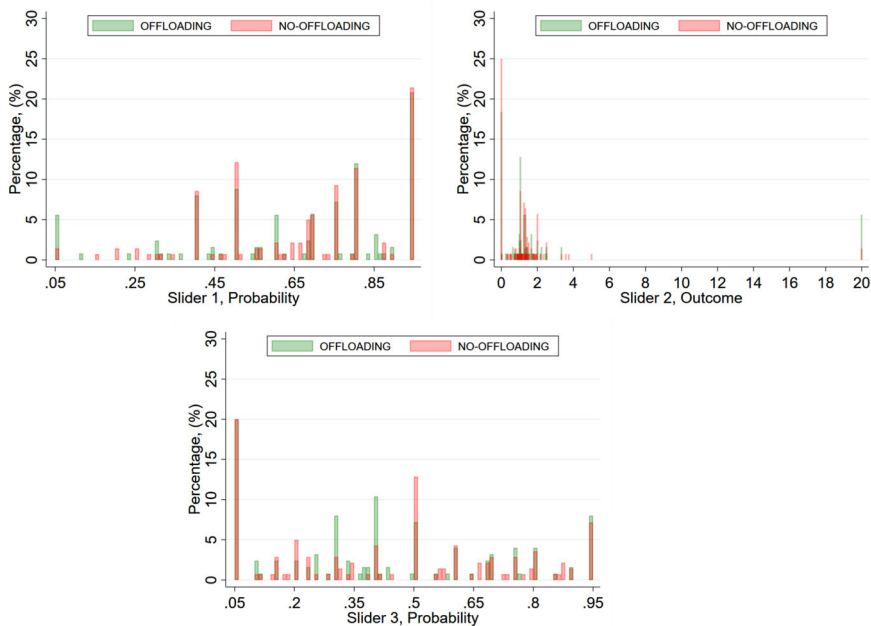


**Fig. 2** Distribution of answers in the first (upper-left), second (upper-right), and third (bottom) budget-choice tasks

### 4.3 Beliefs updating in betting choices

We now turn to the analysis of behavior in the card-betting task. In this part of the experiment participants decided whether to bet or not on risky prospects represented by using an intuitive card game. In particular, the probability of winning the monetary prize in any given trial needed to be estimated by participants based on the provided information (i.e., the remaining winning cards in the deck), hence this is both a task about risky choices as well as a belief updating task. Risk attitudes for the card-betting task are estimated using choices in the lottery task. We remind that in the card-betting task all participants in the OFFLOADING condition were exposed to the offloading tool.

As for the lottery task, we observe participants in the OFFLOADING condition have a higher proportion of “correct” choices (defined as following EV maximization, 62.69%) compared to those in the NO-OFFLOADING condition (60.72%; WMW test,  $N = 293$ .  $z = 2.472$ ,  $p = 0.014$ ). We further study differences in behavior using a random-effect panel probit regressions (Table 2). Model 1 shows that the coefficient of the *Offload* dummy, indicating the treatment effect, is both positive and significant ( $\beta = 0.221$ ,  $SE = 0.050$ ,  $p < 0.001$ ). This can be interpreted as an overall increase in the probability of a expected-value maximizing choice in OFFLOADING compared to the NO-OFFLOADING condition. Model 2 shows the cognitive offloading manipulation is significantly improving the quality of choices

**Table 2** Random-effect probit regression of the probability of a “correct” choice defined as choosing in agreement with EV maximization

| Correct choice  | Model 1             | Model 2             |
|-----------------|---------------------|---------------------|
| Offloading      | 0.221***<br>(0.050) | 0.198***<br>(0.051) |
| Dist_EV         |                     | 0.039<br>(0.042)    |
| Constant        | 0.406***<br>(0.035) | 0.294***<br>(0.110) |
| Controls        | No                  | Yes                 |
| Log L           | −7387.73            | −7017.52            |
| Wald test       | 19.98***            | 29.39***            |
| Obs. (trials)   | 12326               | 11737               |
| Obs. (subjects) | 293                 | 279                 |

Note: † Controls include subject’s Age, Gender, and performance in the Numerical Literacy test. Robust standard errors in brackets, \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

even when controlling for participants’ Age, Gender, and Numerical Literacy, as well as the difficulty of the trial (*Dist\_EV*).<sup>3</sup>

#### 4.3.1 Reliance on offloading tools

In this section we study what drives people to use the provided offloading tools. These are not pre-registered analyses and are hence more exploratory. In previous sections we saw that people scarcely used the offloading tools even if they were free and easy to use. In the lottery task participants used the cognitive offloading tool (i.e., the calculator) 279 times out of the 3744 trials. In particular, only 40 out of 144 (27.8%) participants used the calculator at least once, while for the budge-choice task only three participants used the calculator at least once. Hence, we focus on the lottery task.

The likely predictor of the use of the offloading tools is trial difficulty (in terms of subjective value between the options) and participants’ skills.<sup>4</sup> Among the relevant individual characteristics we measure, numerical literacy (Lipkus et al., 2001) might play a role. This scale consists of three questions related to arithmetic calculations (See Appendix A). It is constructed as the sum of correct choices and it was completed in the absence of any offloading tool for both the OFFLOADING and NO-OFFLOADING groups. Among all 293 participants, 131 achieved perfect numeracy scores (67 in OFFLOADING and 64 in NO-OFFLOADING).

Table 3 shows a random-effects panel probit regression on the probability of using the calculator in the lottery choices. Model 1 shows a negative relation between *EU\_Dist* and the probability of using the calculator ( $\beta = -2.724$ ,  $SE = 0.571$ ,

<sup>3</sup> fourteen participants have missing data for some of the demographic variables.

<sup>4</sup> We remark that all lotteries were computationally very simple, as all had the same number of outcomes, and similar probabilities. Hence, numerical effects are unlikely to play a role.

**Table 3** Random-effect probit regression of the probability of using the offloading aid

| Offloading Usage | Model 1              | Model 2              | Model 3              |
|------------------|----------------------|----------------------|----------------------|
| EU_Dist          | −2.724***<br>(0.571) |                      | −2.538***<br>(0.585) |
| Age              |                      | 0.002<br>(0.016)     | 0.003<br>(0.017)     |
| Female           |                      | 0.140<br>(0.405)     | 0.049<br>(0.423)     |
| Num.Lit          |                      | 0.783**<br>(0.348)   | 0.732**<br>(0.344)   |
| Constant         | −4.487***<br>(0.214) | −5.553***<br>(1.235) | −5.926***<br>(1.220) |
| Log L            | −298.13              | −282.71              | −271.72              |
| Wald test        | 22.77***             | 5.18                 | 22.96***             |
| Obs. (trials)    | 2304                 | 2208                 | 2208                 |
| Obs. (subjects)  | 144                  | 138                  | 138                  |

Note: Robust standard errors in brackets, \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

$p < 0.001$ ). This is to be expected as decisions that are further away from indifference (i.e., for larger expected utility distances) are less difficult and hence would benefit less from the use of the calculator. Model 2 investigates the potential effects of age and gender. These demographic characteristics are not predictive of the use of the offloading aid. Model 3 introduces both choice difficulty (*EU\_Dist*), the demographic factors, and the score in the Numerical Literacy task. We qualitatively reproduce the results of the previous model. Model 2 and 3 also introduce the score participants obtained in the Numerical Literacy task. The latter is indeed significantly correlated with the use of the offloading aid, but in a way which is unexpected. People with higher numerical skills use the calculator more frequently, as indicated by the positive coefficient of the Num.Lit. variable ( $\beta = 0.783$ ,  $SE = 0.348$ ,  $p = 0.024$ ). Under the assumption that people are mostly aware of their numeracy skills, one might expect that the offloading tool should have greater benefits for the group with lower numeracy skills (i.e., who scored lower in the numeracy test) compared to the group that achieved a perfect score. While it is true that the offloading tool, in this case the calculator, is free and easy to use, there is still a “cost” related to using it, in terms of the opportunity cost of time implied in using the provided offloading aid. This is even more salient than in most other studies as we are performing an online experiment where, participants get compensated under an hourly rate (base pay) plus a performance bonus. In this sense, people with high numeracy skills might potentially view the calculator as less desirable as they might be more likely to be able to complete the task without it.

The results regarding the Numerical Literacy task might have an alternative explanation. The score in the Numerical Literacy task might be a proxy for identifying a subset of participants who paid attention to the study more than a classification based on their arithmetic skills. That is, the Numerical Literacy task was not incentivised and

it is likely that participants who still tried to perform well in this part of the study were also those who wanted to perform well overall. In this sense, these participants might also have been more likely to find the calculator less time consuming and hence more appealing to use. This phenomenon might create a mechanical correlation between the score in the Numerical Literacy task and the use of the cognitive offloading tool. While this artifact is not exclusive to our study, or design, and it is likely common to other tasks, it leads us to interpret the reported significant correlation between the use of the offloading tool and the score in the Numerical Literacy task with a grain of salt.

## 5 Conclusions

In an online experiment, where participants were randomly assigned to either receive no help or could use a series of cognitive offloading tools, we show that people's decisions under risk are overall "better" when some of the computations required to arrive at a decision are offloaded to external tools. That is, we observe more choices in alignment with expected value and utility maximization, more consistent choices, and less risk aversion. However, we observe no effect of the cognitive offloading manipulation for what concern the proportion of irrational behaviors, e.g., violations of expected utility (certainty effect), dominated choices, or sub-optimal budget allocations.

Because this was an online experiment, participants who did not get offered the offloading tools could have, in principle, relied on own-sourced aids (e.g., personal calculators). This goes against our hypothesis of finding a difference in decision quality between the two conditions, and hence, while decreasing the probability of finding significant results, it does not hinder our conclusions, but it might partially explain the limited and mixed effects we report in this work.

Arguably the main takeaway of this work is the result that people are not aware of their sub-optimal choices in a way which impedes simple remedies, such as providing them with free, helpful computation offloading tools. In this sense, people are either too confident in the quality of their choices (Dittrich et al., 2005; Berger & Tymula, 2022), are using rules which lead to irrational, sub-optimal behavior (e.g., Bordalo et al., 2012; Herweg & Müller, 2021), or consider the additional time it would take them to use the provided tool not worth the improvement in their decision quality. Quantifying overconfidence in preferential choices, such as risky decisions, is a hard proposition, but some previous works (e.g., Dittrich et al., 2005) showed that there is an over-evaluation of the quality of own (risky) investment decisions in alignment with overconfidence, which is related to task complexity. That is, the more complex is the task, the more overconfident participants are of their own choices. However, Breig and Feldman (2024) showed that participants might not be too confident in their choices over risky prospects. As a matter of fact, when participants are allowed to revise a subset of their initial choices, and told what they previously chose, this leads them to change their mind in a substantial number of times (75%). While this could be attributed to choice inconsistency (Agranov & Ortoleva, 2017; Alós-Ferrer & Garagnani, 2021b), it could also be a way for participants to correct their "mistakes" in favor of more calibrated choices closer to their "true preferences", where the latter is notoriously an elusive concept (Alós-Ferrer et al., 2021a; Alós-Ferrer & Garagnani, 2024; Alós-Ferrer et al.,

2024b). With respect to alternative rules than (subjective) value maximization, several previous works have proposed different theories. Prominently, Bordalo et al. (2012) proposed that a decision maker's attention in risky prospects is drawn to salient payoffs, which leads her to distort true probabilities in favor of the most salient outcomes. In this sense, offloading tools would not be successful in helping overconfident decision makers or those who decided not based on expected (subjective) value maximization, but use alternative rules.

Overall, we show that providing people with tools which help them make better choices under risk generally has a positive but modest effect. The big caveat of this is that people tend to not rely enough on these tools, as they use them parsimoniously, even when they are free. Moreover, irrational behavior is not significantly decreased, which suggests a deeper source of this issue which might not be rooted in a lack of computational capacities. Another interpretation of these results is that there is a systematic discrepancy between normative theories and how people decide, in the sense that people, even when given the chance, might not want to decide in agreement with normative theories.

A final consideration concerns the interpretation of the changes in risk preferences. We find a robust shift in the propensity to accept risk, in the direction of being more risk seeking (i.e., less risk averse), when offloading tools are present. This result should raise some concerns regarding the widespread implementation of cognitive offloading tools as decision aids for important choices. Cognitive offloading tools are most commonly used to help with pension (e.g., Looney & Hardin, 2009) and health choices (e.g., Bright et al., 2012; Moreira et al., 2019). However, if these tools actually shift preferences in a systematic way, especially toward taking more risks, there might be ethical concerns regarding their inattentive implementations, or their effect should at least be considered within the framework of paternalistic interventions.

## Appendix A: Experimental instructions

These are the instructions for each part of the experiment, which were presented separately on screen in Prolific. Text in brackets [...] was not displayed to subjects. In all questions an answer was required before participants were able to proceed.

### [General Instructions]

This study investigates risky decision making in 3 parts and a questionnaire. On top of your fixed earnings, you will earn a bonus payment which will depend on your decisions in the study. Please read all questions carefully. Answer honestly and take care to avoid mistakes. You are expected to earn 5 Pounds per hour. Completing the experiment will take about 30 minutes. The payoffs you see in the experiment will be shown in ECU (Economic Currency Unit), where 10 ECUs equals £1 Pound.

*[Prolific pays in GBP while participants are recruited within Australia]*

### [Explanation of the Lotteries task]

The first task is about choosing between two lotteries. There are 26 rounds in total.

A lottery gives you one of two possible money amounts, each with a certain chance.

For example:

You have a 20% chance of getting 20 ECUs and an 80% chance of getting 10 ECUs. (So, this lottery gives either 20 ECUs with a 20% chance or 10 ECUs with an 80% chance.)

After the study, the computer will randomly pick one of your decisions and play out that lottery. You'll get paid based on the result.

Since any decision could be the one that affects your bonus, it's important to think carefully about each choice.

Note: You are provided an On-Screen Calculator for this task

*[This note is only visible to the OFFLOADING group]*

*[The trial number was visible. Participants needed to click over a lottery to choose it and then confirm their preference. The confirmation button was positioned between the two lotteries to avoid biasing answers based on proximity.]*

### [Explanation of the three Budget-Choice tasks]

In the second task, you'll see three sliders. Each slider lets you pick between two options: a big reward that's less likely to happen or a smaller reward that's more likely to happen.

After you've made your choices, the computer will randomly pick one of the sliders, and the outcome of that slider will determine what you earn. Since any slider could end up being the one that affects your bonus, it's important to think carefully about each decision you make.

Note: You are provided an On-Screen Calculator for this task

*[This note is only visible to the OFFLOADING group]*

*[The trial number was visible. The three sliders used the same graphical representation.]*

### [Explanation of the Beliefs Updating task (Card-Betting Game)]

In the third task, you'll be betting with playing cards over 44 rounds.

You'll see two decks of cards: one red deck and one black deck. Each deck has 10 cards with values from 1 to 10. At the start of each round, one red card and two black cards will be picked at random. Your job is to decide whether or not to bet on the 3rd black card that will be drawn. You win the bet if the 3rd black card is larger than the red card. However, betting costs you something. The costs and rewards (in ECUs) will be shown on the screen.

Here is an example of how a round might look like when the round starts:

*[A graphical example is shown]*

After all the rounds, the computer will randomly pick one round, and the outcome of that round will determine what you earn. Since any round could affect your bonus, it's important to think carefully about each decision you make.

Note: You are provided a Visual Aid (Deck Counter) for this task, it tells you the remaining cards in the black deck. It will show up at the bottom of the screen

*[This note is only visible to the OFFLOADING group]*

*[The trial number was visible]*

### [Questionnaire (Numeracy Literacy)]

*[In the numeracy literacy questionnaire, units are provided in each answer in order to avoid confusion.]*

(Q1) Imagine that we rolled a fair, six-sided die 1,000 times. Out of 1,000 rolls, how many times do you think the die would come up even (2, 4, or 6)?

*[Correct Answer: 500 (times)]*

(Q2) In the BIG BUCKS LOTTERY, the chances of winning a \$10.00 prize is 1%. What is your best guess about how many people would win a \$10.00 prize if 1,000 people each buy a single ticket to BIG BUCKS?

*[Correct Answer: 10 (people)]*

(Q3) In the ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1,000. What percent of tickets to ACME PUBLISHING SWEEPSTAKES win a car?

*[Correct Answer: 0.1 (%)]*

### Screenshots

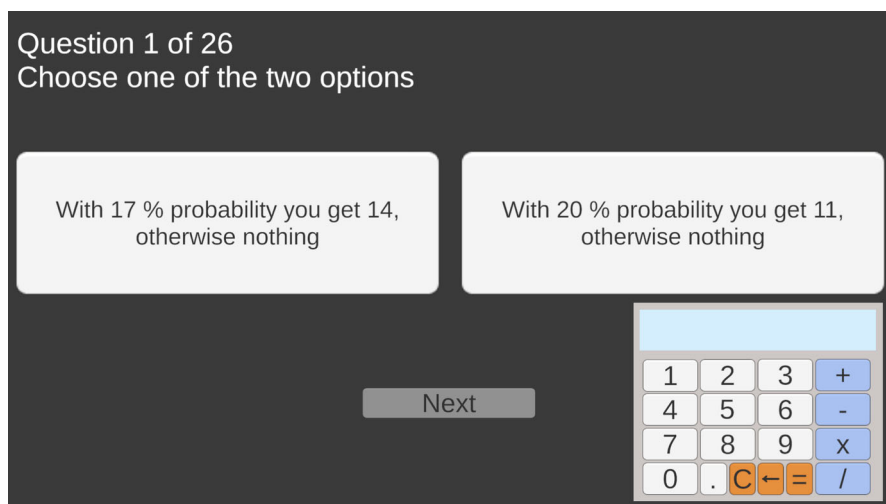


**Fig. 3** Screenshot of the lottery tasks with the OFFLOADING manipulation. NO-OFFLOADING was presented in the same way without the calculator

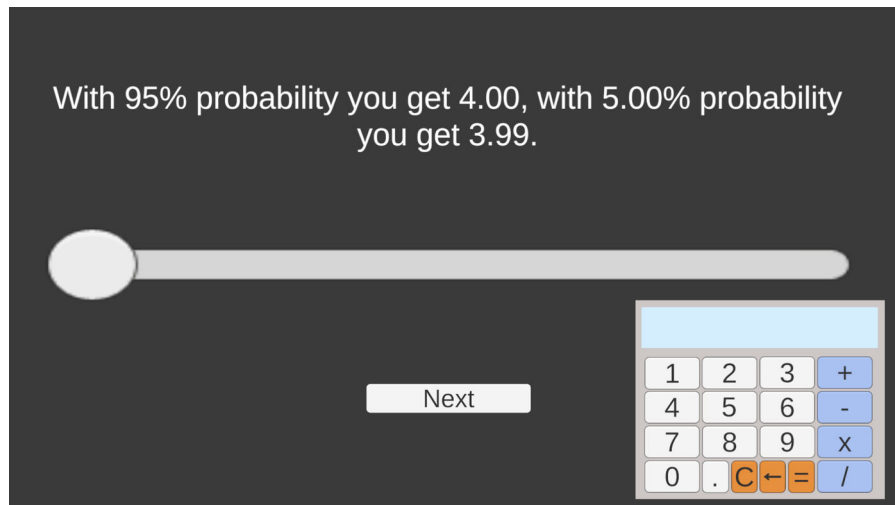


**Fig. 4** Screenshot of the budget-choice tasks with the OFFLOADING manipulation. NO-OFFLOADING was presented in the same way without the calculator



**Fig. 5** Screenshot of the card-betting task with the OFFLOADING manipulation. NO-OFFLOADING was presented in the same way without the last row of cards

## Appendix B: List of lotteries

**Table 4** List of lottery pairs

| Pair | Prob Left | Out Left | Prob Right | Out Right | Type            |
|------|-----------|----------|------------|-----------|-----------------|
| 1    | 80        | 21       | 100        | 16        | Certainty       |
| 2    | 20        | 21       | 25         | 16        | Certainty       |
| 3    | 72        | 27       | 100        | 19        | Certainty       |
| 4    | 18        | 27       | 25         | 19        | Certainty       |
| 5    | 85        | 14       | 100        | 11        | Certainty       |
| 6    | 17        | 14       | 20         | 11        | Certainty       |
| 7    | 52        | 8        | 58         | 10        | Dominated       |
| 8    | 42        | 13       | 36         | 13        | Dominated       |
| 9    | 70        | 7        | 65         | 2         | Dominated       |
| 10   | 40        | 9        | 40         | 11        | Dominated       |
| 11   | 80        | 5        | 75         | 15        | Risk Estimation |
| 12   | 75        | 7        | 65         | 17        | Risk Estimation |
| 13   | 50        | 20       | 70         | 5         | Risk Estimation |
| 14   | 65        | 7        | 50         | 15        | Risk Estimation |
| 15   | 50        | 13       | 60         | 8         | Risk Estimation |
| 16   | 80        | 3        | 40         | 17        | Risk Estimation |
| 17   | 70        | 4        | 35         | 17        | Risk Estimation |
| 18   | 60        | 8        | 35         | 20        | Risk Estimation |
| 19   | 40        | 15       | 75         | 6         | Risk Estimation |
| 20   | 75        | 4        | 35         | 12        | Risk Estimation |
| 21   | 20        | 15       | 55         | 4         | Risk Estimation |
| 22   | 15        | 18       | 65         | 3         | Risk Estimation |
| 23   | 60        | 6        | 35         | 11        | Risk Estimation |
| 24   | 20        | 22       | 80         | 5         | Risk Estimation |
| 25   | 30        | 15       | 75         | 6         | Risk Estimation |
| 26   | 70        | 7        | 50         | 11        | Risk Estimation |

## Appendix C: Restricted-sample analyses

For the lottery task, restricting the sample to include only subjects that used the offloading at least once. Under this restriction we observe an even stronger effect of OFFLOADING (71.09%) relative to NO-OFFLOADING (53.86%; WMW test,  $N = 189$ ,  $z = 3.457$ ,  $p < 0.001$ ). This effect is stronger than the one reported in the main text, suggesting that the actual effect of offloading might be stronger for those who actually relied at least partially on the provided tool. This is also true for expected utility maximization when comparing NO-OFFLOADING to the restricted sample of participants in the OFFLOADING condition which used the calculator at least once

(79.84%; WMW test,  $N = 189$ ,  $z = 4.493$ ,  $p < 0.001$ ). When we restrict the sample to comparing those subjects who used the calculator at least once ( $\mu = 0.643$ ,  $SD = 0.109$ ) to those in the NO-OFFLOADING, we find an even larger difference in the estimated risk attitudes between the two groups (WMW test,  $N = 189$ ,  $z = -7.962$ ,  $p < 0.001$ ).

Table 5 is the equivalent of Table 1 in the main text with the restricted sample of participants who used the offloading tool at least once. The results are stronger but qualitatively the unchanged.

Even in the restricted sample we find not statistically-significant differences in the proportion of dominated choices between treatments (4.38%; WMW test,  $N = 189$ ,  $z = 1.036$ ,  $p = 0.335$ ). A similar (null) result is obtained for the proportion of certainty-effect behavior (51.67%, WMW test,  $N = 189$ ,  $z = 0.509$ ,  $p = 0.607$ ).

Since only three participants used the calculator at least once in this part of the experiment, we refrain from performing the analysis which restricts the sample to those participants who used the calculator.

**Table 5** Random-effect probit regression of the probability of a “correct” choice defined as choosing in agreement with expected value (Models 1) or utility maximization (Models 4 and 52)

| Correct Choice  | Model 1             | Model 2              |
|-----------------|---------------------|----------------------|
| Offloading      | 0.722***<br>(0.236) | 0.483***<br>(0.135)  |
| Age             | 0.000<br>(0.008)    | 0.008*<br>(0.004)    |
| Female          | −0.458**<br>(0.188) | 0.114<br>(0.103)     |
| Num.Lit         | 0.133<br>(0.100)    | −0.053<br>(0.055)    |
| EV_Dist         | 2.745***<br>(0.124) |                      |
| EU_Dist         |                     | 48.603***<br>(8.123) |
| Constant        | −0.979**<br>(0.404) | −0.066<br>(0.165)    |
| Log L           | −1304.07            | −1584.68             |
| Wald test       | 495.38***           | 66.38***             |
| Obs. (trials)   | 2848                | 2848                 |
| Obs. (subjects) | 178                 | 178                  |

Note: The sample includes only those participants who used the cognitive offloading tool at least once.

**Acknowledgements** We gratefully acknowledge financial support from the Center for Brain, Mind, and Market of the University of Melbourne.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions.

**Data Availability** Data is available on OSF at <https://osf.io/sqf3v/>.

## Declarations

**Conflict of interest** The authors have no conflict of interest. No large-language model or similar tools were used in the production of this project.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1), 40–68.
- Allais, M. (1953). Le Comportement de l'Homme Rationnel Devant le Risque, Critique des Postulates et Axiomes de l'Ecole Americaine. *Econometrica*, 21, 503–546.
- Alós-Ferrer, C., Fehr, E., Fehr-Duda, H., & Garagnani M. (2024a). Distinguishing common ratio preferences from common ratio effects using paired valuation tasks: Comment. Working Paper.
- Alós-Ferrer, C., Fehr, E., & Garagnani M. (2024b). Identifying nontransitive preferences. Working Paper, University of Zurich.
- Alós-Ferrer, C., Fehr, E., & Netzer, N. (2021a). Time will tell: Recovering preferences when choices are noisy. *Journal of Political Economy*, 129(6), 1828–1877.
- Alós-Ferrer, C., & Garagnani, M. (2021b). Choice consistency and strength of preference. *Economics Letters*, 198, 109672.
- Alós-Ferrer, C., & Garagnani, M. (2022a). Strength of preference and decisions under risk. *Journal of Risk and Uncertainty*, 64(3), 309–329.
- Alós-Ferrer, C., & Garagnani, M. (2022b). The gradual nature of economic errors. *Journal of Economic Behavior & Organization*, 200, 55–66.
- Alós-Ferrer, C., & Garagnani, M. (2024). Improving risky-choice predictions using response times. *Journal of Political Economy: Microeconomics*, 2(2), 335–354.
- Ballinger, T. P., & Wilcox, N. T. (1997). Decisions, error and heterogeneity. *Economic Journal*, 107(443), 1090–1105.
- Berger, A., & Tymula, A. (2022). Controlling ambiguity: The illusion of control in choice under risk and ambiguity. *Journal of Risk and Uncertainty*, 65(3), 261–284.
- Boldt, A., & Gilbert, S. J. (2019). Confidence guides spontaneous cognitive offloading. *Cognitive Research*, 4(1), 1–16.
- Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Saliency theory of choice under risk. *Quarterly Journal of Economics*, 127(3), 1243–1285.
- Breig, Z., & Feldman, P. (2024). Revealing risky mistakes through revisions. *Journal of Risk and Uncertainty*, 68(3), 227–254.
- Bright, T. J., Wong, A., Dhurjati, R., Bristow, E., Bastian, L., Coeytaux, R. R., Samsa, G., Hasselblad, V., Williams, J. W., Musty, M. D., et al. (2012). Effect of clinical decision-support systems: A systematic review. *Annals of Internal Medicine*, 157(1), 29–43.

- Burnett, L. K., & Richmond, L. L. (2023). Just write it down: Similarity in the benefit from cognitive offloading in young and older adults. *Memory & Cognition*, 51(7), 1580–1592.
- Cettolin, E., & Riedl, A. (2019). Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization. *Journal of Economic Theory*, 181, 547–585.
- Choi, S., Fisman, R., Gale, D., & Kariv, S. (2007). Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review*, 97(5), 1921–1938.
- Choi, S., Kariv, S., Müller, W., & Silverman, D. (2014). Who is (more) rational? *American Economic Review*, 104(6), 1518–1550.
- Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, 140(1), 102–116.
- Costa, A. J., Silva, J. B. L., Chagas, P. P., Krinzinger, H., Lonneman, J., Willmes, K., Wood, G., & Haase, V. G. (2011). A hand full of numbers: A role for offloading in arithmetics learning? *Frontiers in Psychology*, 2, 368.
- Cubitt, R. P., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, 1(2), 115–131.
- Dittrich, D. A., Güth, W., & Maciejovsky, B. (2005). Overconfidence in investment decisions: An experimental approach. *The European Journal of Finance*, 11(6), 471–491.
- Dumm, R. E., Eckles, D. L., Nyce, C., & Volkman-Wise, J. (2020). The representative heuristic and catastrophe-related risk behaviors. *Journal of Risk and Uncertainty*, 60(2), 157–185.
- Engelmann, J. B., Capra, C. M., Noussair, C., & Berns, G. S. (2009). Expert financial advice neurobiologically “Offloads” financial decision-making under risk. *PLoS ONE*, 4(3), e4957.
- Garagnani, M. (2023). The predictive power of risk elicitation tasks. *Journal of Risk and Uncertainty*, 67(2), 165–192.
- Garagnani, M., & Vieider, M. F. (2002). Economic consequences of numerical adaptation. Working Paper.
- Gillenkirch, R. M., & Velthuis, L. (2023). Delegated risk-taking, accountability, and outcome bias. *Journal of Risk and Uncertainty*, 67(2), 137–161.
- Glimcher, P. W. (2022). Efficiently irrational: Deciphering the riddle of human choice. *Trends in Cognitive Sciences*, 26(8), 669–687.
- Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, 112(2), 631–645.
- Halevy, Y., Persitz, D., & Zrill, L. (2018). Parametric recoverability of preferences. *Journal of Political Economy*, 126(4), 1558–1593.
- Halevy, Y., Walker-Jones, D., & Zrill, L. (2023). *Difficult Decisions*. Department of Economics: University of Toronto.
- Herweg, F., & Müller, D. (2021). A comparison of regret theory and salience theory for decisions under risk. *Journal of Economic Theory*, 193, 105226.
- Hey, J. D., Morone, A., & Schmidt, U. (2009). Noise and bias in eliciting preferences. *Journal of Risk and Uncertainty*, 39(3), 213–235.
- Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, 62(6), 1291–1326.
- Hey, J. D., & Pace, N. (2014). The explanatory and predictive power of non two-stage-probability theories of decision making under ambiguity. *Journal of Risk and Uncertainty*, 49(1), 1–29.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- Kelman, M., Fallas, D., & Folger, H. (1998). Decomposing hindsight bias. *Journal of Risk and Uncertainty*, 16(3), 251–269.
- Kurtz-David, V., Persitz, D., Webb, R., & Levy, D. J. (2019). The neural computation of inconsistent choice behavior. *Nature Communications*, 10(1), 1583.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21(1), 37–44.
- Loomes, G., Moffatt, P. G., & Sugden, R. (2002). A microeconomic test of alternative stochastic theories of risky choice. *Journal of Risk and Uncertainty*, 24(2), 103–130.
- Looney, C. A., & Hardin, A. M. (2009). Decision support for retirement portfolio management: Overcoming myopic loss aversion via technology design. *Management Science*, 55(10), 1688–1703.
- Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley.
- McFadden, D. L. (2001). Economic choices. *American Economic Review*, 91(3), 351–378.

- McGranaghan, C., Nielsen, K., O'Donoghue, T., Somerville, J., & Sprenger, C. D. (2024). Distinguishing common ratio preferences from common ratio effects using paired valuation tasks. *American Economic Review*, 114(2), 307–347.
- Moffatt, P. G. (2015). *Experiments for experimental economics*. Palgrave Macmillan.
- Moreira, M. W., Rodrigues, J. J., Korotaev, V., Al-Muhtadi, J., & Kumar, N. (2019). A comprehensive review on smart decision support systems for health care. *IEEE Systems Journal*, 13(3), 3536–3545.
- Morrison, A. B., & Richmond, L. L. (2020). Offloading items from memory: Individual differences in cognitive offloading in a short-term memory task. *Cognitive Research: Principles and Implications*, 5(1), 1.
- Risko, E. F., & Dunn, T. L. (2015). Storing information in-the-world: Metacognition and cognitive offloading in a short-term memory task. *Consciousness and Cognition*, 36, 61–74.
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688.
- Ruggeri, K., Alf, S., Berge, M. L., Bertoldo, G., Bjørndal, L. D., Cortijos-Bernabeu, A., Davison, C., Demić, E., Esteban-Serna, C., Friedemann, M., et al. (2020). Replicating patterns of prospect theory for decision under risk. *Nature Human Behaviour*, 4(6), 622–633.
- Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333(6043), 776–778.
- Storm, B. C., & Stone, S. M. (2015). Saving-enhanced memory: The benefits of saving on the learning and remembering of new information. *Psychological Science*, 26(2), 182–188.
- Taylor, M. P. (2013). Bias and brains: Risk aversion and cognitive ability across real and hypothetical settings. *Journal of Risk and Uncertainty*, 46(3), 299–320.
- Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review*, 34(4), 273–286.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76(1), 31–48.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.