Risky Choice and Memory for Effort: Hard Work Stands Out

Mason, A.[1], Madan, C.R.[2], Freas, C.A.[3], Simonsen, N.[2], Ludvig, E.A.[1], Spetch, M.L.[3]

1. Department of Psychology, University of Warwick
2. School of Psychology, University of Nottingham
3. Department of Psychology, University of Alberta

Author Note

The datasets generated during the current study are available on OSF https://osf.io/695js/, along with the code used to run the experimental tasks.

Abstract

When deciding between different courses of action, both the potential outcomes and the costs of making a choice should be considered. These costs include the cognitive and physical effort of the different options. In many decision contexts, the outcome of the choice is guaranteed but the amount of effort required to achieve that outcome is unknown. Here we studied choices between options that varied in the riskiness of the effort (number of responses) required. People made repeated choices between pairs of options that required them to click different numbers of sequentially presented response circles. Easy-effort options led to small numbers of response circles, whereas hard-effort options led to larger numbers of response circles. For both easy- and hard-effort options, fixed options led to a consistent effort, whereas risky options led to variable effort that, with a 50/50 chance, required more or less effort than the fixed option. Participants who showed a preference for easier over harder options were more risk averse for decisions involving hard options than for decisions involving easy options. On subsequent memory tests, people most readily recalled the hardest outcome, and they overestimated its frequency of occurrence. Memory for the effort associated with each risky option strongly correlated with individual risky preferences for both easy-effort and hard-effort choices. These results suggest a relationship between memory biases and risky choice for effort similar to that found in risky choice for reward. With effort, the hardest work seems to particularly stand out.

Imagine that you are cooking dinner and you realise you are missing a key ingredient. You could walk over to the supermarket that always stocks the ingredients, or you could walk a shorter distance in the opposite direction to the shop that may or may not have the ingredient, risking that you may still have to walk to supermarket. Whether or not you get the ingredients is not the focus of your decision, instead you are weighing up how much effort to exert to get them. Understanding risky choice has been the focus of extensive research in psychology and behavioral economics (e.g., Hertwig & Erev, 2009; Kahneman & Tversky, 1979), as well as in various other disciplines such as biology (e.g., Kacelnik & Bateson, 1996), medicine (e.g., Reyna & Lloyd, 2006; Simianu et al., 2016), neuroscience (e.g., Platt & Huettel, 2008), and politics (e.g., Vis & Van Kersbergen, 2007). Most experimental studies of risky choice in humans, however, have focused on choices between options that differ in the risk associated with the outcome value (e.g., risky or fixed amounts of monetary rewards). In the scenario outlined above, you are choosing between two options that provide the same eventual reward, but one option involves a fixed amount of effort, and the other option involves risk in which you might save some effort or you might end up exerting even more effort. In the present research, we aimed to extend the study of risky choice in humans to situations where the risk involved the cost (i.e., effort) needed to obtain an outcome, a key component of many everyday choices.

From both a biological and behavioral perspective, effort should be a salient determinant of choice. For example, when foraging, how much time and energy is expended to obtain these nutrients can be as important as the nutrient obtained (e.g., Charnov, 1976). For economic decisions, the costs, which can include money, time, and physical or cognitive effort can be as important as the benefits (e.g., Kool, McGuire, Rosen & Botvinick, 2010; Otto & Daw, 2019). Indeed, the role of effort in choice has been the focus of an increasing number of studies, and it

has been argued that the work required to obtain a reward is a critical determinant of behavior and should "receive its own spotlight" (Salamone, Correa, Yang, Rotolo & Presby, 2018, p.2).

To date, most studies on effort-based choice behavior have focused on how effort affects decisions between options that provide different rewards, or on how effort and reward trade off in determining choice. For example, increases in effort increase preference for a small, certain reward over a larger, uncertain reward, both in a risk-sensitive foraging task with rats (Kirshenbaum, Szalda-Petree & Haddad, 2000) and in marketing research with humans (Kivetz, 2003). In humans, increases in effort (via difficulty of mathematical calculations) enhanced brain sensitivity to the magnitude of rewards and losses (Hernandez et al., 2013). The value attached to monetary reward decreases with greater effort required to obtain it, known as effort discounting (e.g., Botvinick, Huffstetler, & McGuire, 2009; Hartmann, Hager, Tobler & Kaiser, 2013). In the brain, dopamine plays a role in choices involving trade-offs between effort and reward amount in both humans (e.g., Treadway, Bossaller, Shelton & Zald, 2012) and non-human animals (see Salamone et al., 2018 for a recent review). Effort and amount are processed via different neural pathways (the cingulate cortex and ventromedial prefrontal cortex, respectively) before being integrated for decisions involving effort-reward trade-offs (Klein-Flügge, Kennerley, Friston & Bestmann, 2016).

The importance of effort in human decision making is underscored by evidence that deficits in effort-based decision making, characterized by less willingness to exert effort for a higher reward amount, have been implicated in schizophrenia (e.g., Gold et al., 2013) and depression (e.g., Treadway, Bossaller, et al, 2012). Moreover, an effort-reward imbalance has been identified as an important factor in workplace stress (Eddy et al., 2016), and a recent study with teenagers found lower sensitivity to effort costs in adolescents than in adults (Sullivan-

Toole, DePasque, Holt-Gosselin & Galván, 2019). Despite the considerable research on how effort and reward trade off in risky choice (Otto & Daw, 2019), much less is known about how people choose between options that provide the same rewards and differ only in the riskiness of the effort involved.

When rewards differ in magnitude, risky choice depends on the set of outcomes in the decision context (see Madan, Spetch, Machado, Mason & Ludvig, 2021). When monetary outcomes are learned through experience, people often show context-dependent biases in which they are more risk seeking for choices involving the best outcomes in the context (e.g., gains or high-value rewards) than for choices involving losses or lower-value rewards (e.g., Konstantinidis, Taylor & Newell, 2018; Ludvig, Madan & Spetch, 2014). This pattern of results is opposite to that seen in decisions from description (Kahneman & Tversky, 1979; Ludvig & Spetch, 2011), and appears to reflect overweighting of the extreme (best and worst) outcomes in memory (Madan, Ludvig & Spetch, 2014). Post-choice memory tests showed that people were more likely to recall the best and worst outcomes and to report that they occurred more often than the intermediate outcomes, and memory biases correlated with individual levels of risk preference (Madan, Ludvig & Spetch, 2017).

Only a small number of studies have investigated how people choose between fixed and risky effort when reward is held constant (Apps, Grima, Manohar & Husain, 2015; Meyer, Schley & Fantino, 2011; Nagengast, Braun & Wolpert, 2011), and none of these focused on context-dependent biases in risky choice or memory. Here we tested whether people would show biases for risky effort that align with those seen for risky reward. If so, people would be more risk seeking for choices involving easy-effort outcomes (i.e., the better outcomes) than for choices involving hard-effort outcomes in an experience-based task. We also tested whether

people would show memory biases for the easiest and hardest effort levels and whether biases in memory for effort would correlate with individual levels of risky choice.

The trade-off between effort and reward amount suggests that effort-based choice may show similar biases to reward-based choice. Effort costs, however, may sometimes have different qualities than reward costs. Although people and animals usually choose to minimize the time and effort required to obtain a goal, increased effort sometimes leads to increases in the subjective value of the outcome obtained, and in some cases, organisms will paradoxically choose options that require more effort (Inzlicht, Shenhav & Olivola, 2018; Kacelnik & Marsh, 2002; Zentall, 2010). Several species, including humans, sometimes show "contrafreeloading", choosing to work for reward over receiving it for free (e.g., Jensen, 1963; Navarro & Osiurak, 2015; Osborne, 1977; Rosenberger, Simmler, Nawroth, Langbein & Keil, 2020; Tarte, 1981). For example, people will pay money to exert physical effort at a gym, and the popularity of puzzles and sudoku suggest that people will choose to exert cognitive effort in the absence of any monetary reward. Because of these paradoxical findings, it remains unclear whether decisions involving risky effort would show risk preferences and biases similar to those that have been reported for decisions involving risky rewards.

Here we sought to examine how people respond to risk in effort level in the absence of differential rewards. A set of 3 experiments examined how people choose between fixed and risky effort, and how they remember the effort levels they experience. The experiments also contribute to the literature on risky decision making by assessing whether known biases in risky choice and memory for rewards generalize to choices based on effort. In the experiments, participants made repeated experience-based choices between options that differed in the level and variability of effort (number of spatially distributed mouse clicks) required to complete the

trial. Two options were "easy", requiring only a few responses, whereas the other two options were "hard" and required more responses. One easy and one hard option were "fixed", such that the required number of responses was the same every time that option was chosen. The other two options were risky, sometimes requiring more and sometimes requiring fewer responses than the corresponding fixed options. Table 1 shows the effort levels for each option. Choices between easy and hard options assessed effort preference, and choices between fixed and risky options assessed risk preference. Participants were given the same monetary reward after completing all trials regardless of which options they chose. After completing a series of choice trials, we tested participants' memory of the effort associated with each risky option.

Experiment 1 used in-person testing, and Experiments 2 and 3 were conducted using the online platform Prolific Academic. Experiment 3 controlled the time taken to complete the effort requirement to disentangle the effects of effort and time. All data, materials and pre-registration documents are available on the Open Science Framework [https://osf.io/695js/]).[1]

[1] In addition to the reported experiments, we also conducted two aborted studies and one additional study reported in supplementary materials. The first aborted study was conducted prior to Experiment 1 and was aborted because comments made by participants suggested, and an examination of the data confirmed, that most participants were not learning which were the easier options. We therefore increased the response requirement for the harder effort options and started the current Experiment 1. Another experiment was initiated prior to Experiment 2 but was aborted early because in-person testing was no longer possible due to Covid-19. The experiment reported in supplemental materials was conducted prior to Experiment 3 and was our first attempt to control time across effort levels. For that study, many participants failed to complete the effort requirement within the specified time limit on a substantial number of trials, making the results inconclusive. We therefore adjusted the time limits and effort levels and repeated the experiment, reported here as Experiment 3.

Table 1. Number of required responses (circles to click) for each choice option.

	Easy Fixed	Easy Risky	Hard Fixed	Hard Risky	Time controlled
Experiment 1	3	1 or 5	9	7 or 11	No
Experiment 2	2	1 or 3	8	7 or 9	No
Experiment 3	3	1 or 5	9	7 or 11	Yes

Experiment 1

In this experiment, participants chose between pairs of doors that led to different numbers of responses required to end the trial. An easy-fixed door required 3 responses, an easy-risky door required 1 or 5 responses with equal probability, a hard-fixed door required 9 responses and a hard-risky door required 7 or 11 responses with equal probability (see Table 1). Based on how people respond to experienced outcomes in risky choice (e.g., Ludvig et al, 2014; Madan et al., 2014), we expected that people would overweight the hardest effort level (11 responses) and easiest effort level (1 response) in both choice and memory. Accordingly, we pre-registered one primary hypothesis about choice and three secondary hypotheses about memory. The primary hypothesis was that people would overweight the hardest effort option and therefore make fewer risky choices for decisions between hard options than for decisions between easy options. The secondary hypotheses were that (1) people will be more likely to report extreme numbers of responses (1 and 11) on the first-outcome recall test, (2) people will overestimate the frequency of these extreme numbers of responses (1 and 11), relative to the equally often experienced non-

extreme numbers (5 and 7), and (3) individuals' responses on the first-recall and the frequency-judgements tests will correlate with their risky choices.

Methods

Participants

We recruited 104 participants (54 Male, 50 Female; age range of 18 to 26 with mean age of 19) from the University of Alberta Psychology participant pool. Participants earned course credit and were paid $5 (Canadian) as a bonus for completing the experiment. They were informed that they needed to complete 200 choice trials and answer a few memory questions in order to obtain the $5 bonus. All participants provided informed consent, and ethics approval was provided the University of Alberta Human Research Ethics Board.

Procedure

Up to 15 participants signed up for each time slot, and they first sat as a group in a central room to receive general instructions and provide written informed consent. They were then assigned to individual testing rooms, where they individually completed the task on PC computers running Windows 10 and using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA).

At the beginning of each trial, participants were shown pictures of one or two visually distinct doors (Figure 1A). Clicking a door with the mouse was immediately followed by removal of the door image(s) followed by the sequential presentation of one or more black response circles, with the number of circles dependent on which door was clicked (Figure 1B). Response circles were presented one at a time in locations randomly selected (with replacement) from 9 evenly spaced locations on the computer screen. A 500-ms delay preceded each

presentation of a response circle, and the circle remained on the screen until it was clicked with the mouse. The mouse cursor reset to the middle of the screen before each response circle was presented. After the last circle for the trial was clicked, a trial counter displayed at the bottom of the screen incremented by one count and the next trial began (Figure 1B).

Figure 1A shows the four door images used in the experiment and the contingencies between these four choice options and the six numbers of response circles. The door image assigned to each choice option was counterbalanced across participants, and the left-right location of each door was counterbalanced across trials within blocks. The easy-fixed door was always followed by 3 response circles whereas the easy-risky door was followed by a 50/50 chance of either 1 or 5 response circles. The hard-fixed door was always followed by 9 response circles, whereas the hard-risky door was followed with a 50/50 chance of either 7 or 11 response circles.

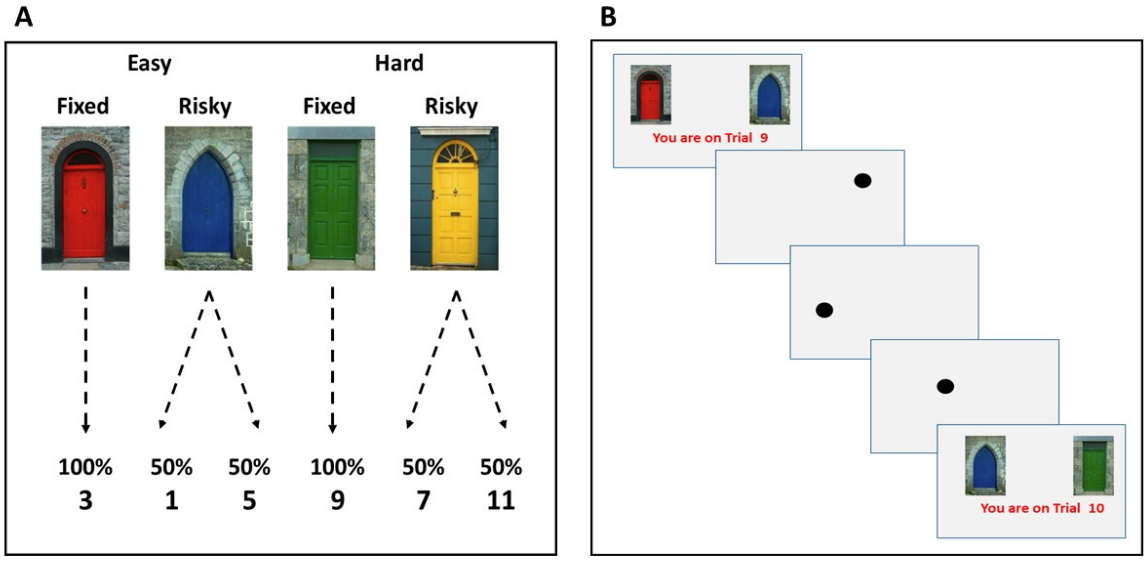

Figure 1. (A) Schematic illustrating the choice stimuli and effort contingencies in Experiment 1. The numbers indicate how many response circles needed to be clicked to complete the trial. Fixed doors led to the same number each time (100%) whereas risky

doors led equally often (50%) to two different numbers. The specific doors associated with each effort contingency were counterbalanced across participants. (B) Schematic of an example choice trial in which the easy-fixed door was selected and was followed by 3 response circles. Participants needed to click on one of the doors to choose it and then needed to click on each of the successively presented response circles to complete the trial. A 500-ms delay preceded the presentation of each response circle. The images shown are not exactly to scale.

During the choice phase of the experiment, participants were presented with three types of trials. *Single-option trials* presented only a single door that the participants were required to click to continue. These trials ensured that the participants experienced the effort levels associated with each door throughout the experiment regardless of their choices. *Effort-preference trials* presented a choice between an easy-fixed door and a hard-fixed door, or between an easy-risky door and a hard-risky door, i.e., objectively different effort levels that did not differ in risk. These trials assessed whether participants had learned the door-effort contingencies and were choosing to minimize effort. As per the pre-registration on OSF, and consistent with the criterion used in previous studies (e.g., Ludvig & Spetch, 2011), only participants who chose the easy options on 60% or more of the effort-preference trials were included in the primary analyses. This criterion excludes participants who failed to learn the task contingencies or were not motivated to minimize effort and chose randomly (Ludvig & Spetch, 2011). With 80 total effort-preference trials, 48 low-effort responses (60%) represent the lowest number that is reliably different from random responding (at p =0.05, using cumulative binomial probability). Finally, *risk-preference trials* provided a choice between an easy-fixed and an easy-

risky door, or between a hard-fixed and a hard-risky door. These risk-preference trials provided choices between doors that required the same average effort, but one was fixed and one was risky. Thus, these trials provided a measure of risk preference for each level of effort.

The choice phase consisted of five blocks of trials, separated by a brief break (an on-screen riddle). The right and left location of each door was counterbalanced for each trial type in each block. Each block provided 8 single-choice trials (two for each door), 16 effort-preference trials (4 for each easy and hard door combination), and 16 risk-preference trials (8 easy-effort decisions and 8 hard-effort decisions), making 40 trials per block, and 200 trials in total.

Following the choice phase, participants were given two types of memory tests. First, they were given a *First-Recall* test in which each of the four doors was presented one at a time (in random order for each participant); for each door, the participant was instructed on the screen to type the first number of response circles that came to mind. This test was designed to assess how accessible each response number was in the participant's memory. The tests assumes that even if both outcomes following a risky door can be recalled, there may availability biases in that one of the outcomes may come to mind quicker than the other one. Second, they were given a *Remembered-Frequency* test, in which they were again shown each door, in a new randomly determined order, and below the door they saw six numbers corresponding to the six numbers of response circles (i.e., 1, 3, 5, 7, 9, 11) experienced in the task. The participant was instructed on the screen to type the percentage of time they had encountered each number of response circles following the displayed door.

Results

Only 65 of the 104 participants passed the criterion of choosing the easy options on 60% or more of the effort-preference trials on the last two blocks, and as per the pre-registration, only

the data from these 65 participants were used in the analyses reported below. Of the participants who did not meet criterion, 15 chose the hard option on 60% or more of the effort-preference trials. These high-effort choosers spent an average of 6.7 minutes longer on the choice task than the low-effort choosers, highlighting the cost of choosing high-effort options. Exploratory analyses on the 15 high-effort choosers are reported in the supplemental materials.

As per the pre-registration, all t-tests were one-tailed. As shown in Figure 2A, people developed risk aversion for decisions involving hard options but not for decisions involving easy options. Averaged over the last two blocks (Figure 2B), participants chose the risky option 16.0 ± 4.9 percentage points less often for choices involving hard options (32.8 ± 3.7%) than for choices involving easy options (48.8 ± 4.5%), $t(64) = 3.26$, $p = .002$, $d = .40$.

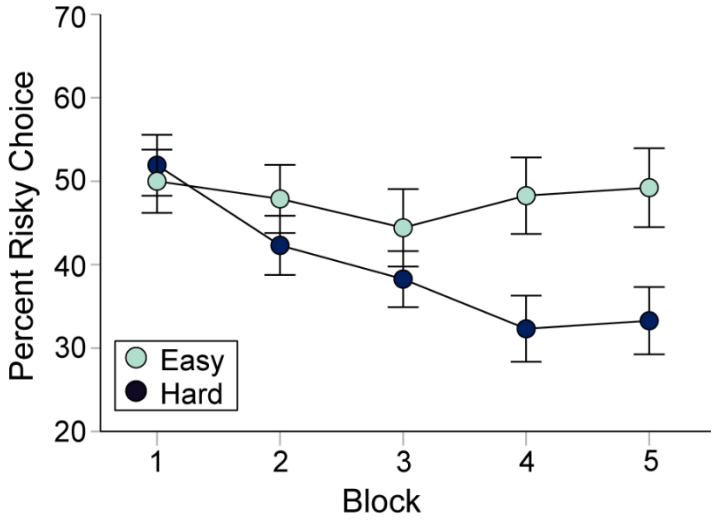

Figure 2. Risky choice results for Experiment 1. Mean percentage (± SEM) of risky choices for the decisions involving easy or hard options for each block of choice trials.

For the memory results, participants were only included in each analysis if they had provided a valid response for the relevant memory test. On the first-recall test, participants showed a bias toward reporting the hardest response requirement. Figure 3A shows the percentage of participants who reported 1 or 5 for the easy-risky door and 7 or 11 for the hard-risky doors. For the easy-risky door, there was no difference between the percentage of participants who reported 1 or 5, χ^2 (1, N = 54) = 0, p = 1. For the hard-risky door, however, more participants reported the high-extreme number (11) than the non-extreme number (7), χ^2 (1, N = 42) = 4.67, p = .031.

Although group-level biases in the recall test appeared only for the hard-risky door, responses on this memory test correlated with individuals' choice behavior for both risky doors. Figure 3B plots risk preference in the choice task on the basis of responses on the first-recall test. For the easy-effort choices, people who recalled 1 response showed a higher percentage of risky choices (62.3 ± 7.0%; N = 28) than those who recalled 5 responses (30.4 ± 4.9%; N = 28), t(54) = 3.72, p < .001, d =0.99. Similarly, for the hard-effort choice, people who recalled 7 responses showed a higher percentage of risky choices (59.4 ± 6.5%; N = 14) than those who recalled 11 responses (12.5 ± 4.0%; N = 28), t(40) = 6.47, p < .001, d =2.12. To factor out the contribution of any differences between people in their frequency of experiencing each outcome, we conducted a partial correlation between risky choice and the recalled number for each risky choice, with obtained frequency of each outcome as the controlled variable (see Madan et al., 2014, 2017). This partial correlation was significant, even when the obtained frequency of each outcome for each risky door was controlled (easy: r_p (53) = -.44, p = .001; hard: r_p (39) = -.68, p < .001).

On the remembered-frequency test, participants showed a bias in reporting the effort frequency for the hard-risky door but not for the easy-risky door. Figure 3C shows the mean reported frequency (in percent of trials) of 1 or 5 responses for the easy-risky door and of 7 or 11 responses for the hard-risky door. For the easy-risky door, participants did not report a higher frequency of occurrence for the extreme (1) number of responses than for the non-extreme (5) number of responses, $t(60) = 0.04$, $p = 1.0$, $d = 0.005$. For the hard-risky door, however, participants reported the extreme number (11) of responses as having occurred 25.1 ± 6.3 percentage points more often than the non-extreme number (7) of responses, $t(62) = 4.00$, $p < .001$, $d = 0.50$.

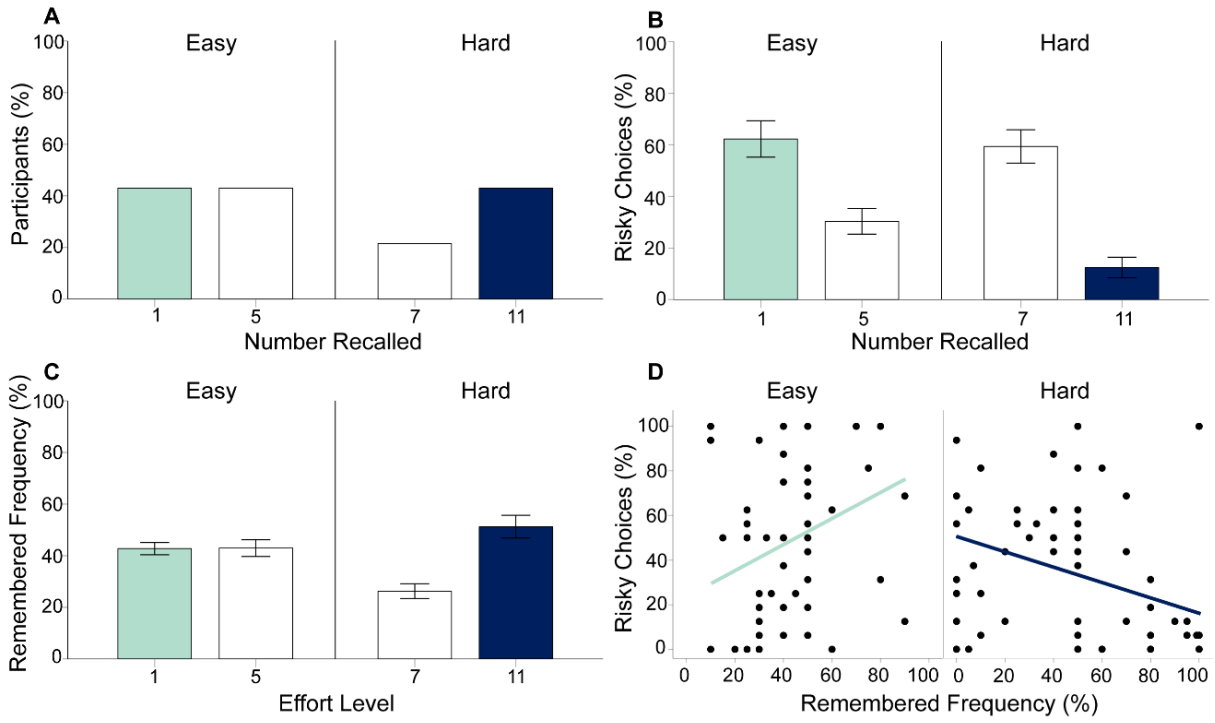

Figure 3. Results of the memory tests and correlations with risky choice in Experiment 1. (A) Percentage of participants who responded with 1 or 5 for the easy-risky door, and with

7 or 11 for the hard-risky door on the first-recall test. (B) Mean risk preference (±SEM) for easy-effort and hard-effort choices, split by answer on the first-recall test. (C) Mean percentage (±SEM) reported on the remembered frequency test that 1 or 5 response circles occurred on the easy-risky door and that 7 or 11 response circles occurred on the hard-risky door. (D) Scatterplot of risk preference on easy-effort decisions as a function of remembered frequency of the easiest outcome (1 response) and risk preference on hard-effort decisions as a function of remembered frequency of the hardest outcome (11 responses). Each dot represents an individual participant, and the lines indicate the linear regression.

Figure 3D plots risk preference in the last 2 blocks against remembered frequency of the extremes (1 or 11 responses). For the easy-effort decisions, risky choices increased with judged frequency of the easy extreme (1 response), r (59) = .30, p = .020. even when controlling for outcomes experienced, r_p (58) = .28, p = .028. For the hard-effort decisions, risky choices decreased with judged frequency of the hardest extreme (11 responses), r (61) = -.39, p = .001. even when controlling for outcomes experienced, r_p (60) = -.35, p = .006. Thus, individual differences in the remembered frequency of the different amounts of effort correlated significantly with risky choice for decisions involving both easy and hard options.

Experiment 2

This study provided a replication and extension of Experiment 1 using a larger sample of participants recruited from Prolific Academic and with some variations in the procedure. Because so many participants in Experiment 1 did not develop a strong preference for the easy

options, we made several procedural changes designed to facilitate learning of the effort level associated with each choice door: (1) indicating the number of required responses immediately after selection of a door, (2) inserting a delay between each response to make the differences in effort more salient, and (3) using a new set of response numbers (as shown in Table 1) to make the easy and hard sets more distinct. For participants who chose easy options on effort-preference trials, our pre-registered predictions were that they would choose the risky option more often on decisions involving easy options than on decisions involving hard options and they would be more likely to report the easiest and hardest outcomes than intermediate outcomes on a recall test.

For this study, we also used visually distinct response circles that were consistently paired throughout the session with the number of responses (1, 2, 3, 7, 8, or 9) required to complete the trial as shown in Figure 4. The purpose of this variation was to determine whether we could identify and characterize a subset of people who show a paradoxical preference for high effort (e.g., Inzlicht et al., 2018). Specifically, if some individuals consistently choose harder options, these individuals may show opposite patterns of risky choice than those who prefer easy options, and they may show a preference for stimuli associated with the high effort (similar to the "IKEA effect", Norton, Mochon & Ariely 2012). Because very few participants chose high-effort options in this experiment, however, we had insufficient power to address these questions and therefore all analyses related to the stimulus preferences are reported in supplemental materials.

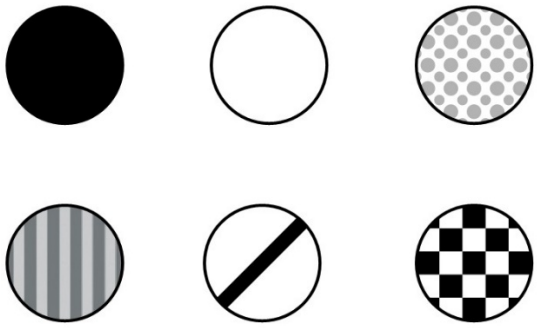

Figure 4. Images of circle patterns associated with the number of responses (1, 2, 3, 7, 8, 9) required to complete the trial in Experiment 2. The number of responses was randomly assigned to each circle pattern for each participant.

Methods

Participants

We recruited 250 participants from Prolific Academic. Participants were paid £7 for completing the experiment. They were informed that they needed to complete 128 choice trials plus some memory and preference tests in order to earn their completion code and that the task should take approximately 45 min to complete. Thirteen participants were excluded because they were either not recorded on Prolific (N=1), exceeded the Prolific time limit of 115 min (N=1) or restarted the experiment after completing some trials (N=11). These exclusions left 237 participants (154 males, 80 females, age range of 18 to 65 with mean of 27).

Procedure

The program was created in PsychoJS and run on the Pavlovia platform (Peirce et al., 2019). The procedure was the same as that used in Experiment 1 with the following exceptions: Clicking on a choice door was followed by a 2-s message that stated “You will need to click

[number] circle[s]", with the number being in the set 1,2,3,7,8,9 and determined by which door was clicked. Each required number of responses was associated with a different visual pattern on the response circles. A 1.5-s delay with a blank screen preceded the presentation of each sequentially presented response circle, and a 3-s delay with a blank screen preceded the onset of each new trial. As this experiment was run online the mouse was not re-centered between trials. There was no trial counter display, but at the end of Blocks 2 and 3 a message indicated the number of trials completed thus far. The door images assigned to each choice option and the circle patterns assigned to each effort level were randomly assigned for each participant. The number of required responses for each door was as follows: easy-fixed door = 2, easy-risky door = 1 or 3 with a 50/50 chance, hard-fixed door = 8, and hard-risky door = 7 or 9 with a 50/50 chance. The session included 128 choice trials divided into four blocks. The first block was a short learning block and consisted of eight single-option trials, two with each door presented alone, counterbalanced across door location. Each risky door provided one instance of each of its response requirements during the learning block. The next three blocks each included eight single-option trials (two for each door), 16 effort-preference trials (eight with risky options and eight with fixed options), and 16 risk-preference trials (eight with easy options and eight with hard options) for a total of 40 trials per block. All trial types were counterbalanced for side.

After the last block of choices, all participants were given a *First-Recall* test like that described in Experiment 1 in which participants were asked to type the first number of response circles that came to mind for each door. This test was followed by two tests about the circle patterns that are described in the supplemental materials.

Results

As per the pre-registration, we used choices on effort-preference trials to partition the set of participants into low-effort choosers (chose easy doors on 60% or more of the effort-preference trials) or high-effort choosers (chose hard doors on 60% or more of the effort-preference trials). Because there were only 3 full choice blocks, we used the results from the last block of choice trials (i.e., after learning occurred) for effort-preference and risk-preference analyses. In this experiment (unlike Experiment 1), a large majority of participants chose the easy doors, and hence this partitioning led to 219 low-effort choosers and only six high-effort choosers. Results for the six participants who chose hard doors on effort-preference trials are presented in the Supplemental Materials. The results reported below are for the 219 participants who chose easy options on effort-preference trials. All *t*-tests reported are one-sided.

People were more risk averse for decisions involving hard options than for those involving easy options, consistent with Experiment 1. Figure 5A shows the percentage of risky choices made when participants chose between easy doors or between hard doors across blocks of choices. On the last block, participants chose the risky option 9.0 ± 2.7 percentage points less often for the hard-effort decision (38.9 ± 2.1%) than for the easy-effort decision (47.8 ± 2.4%), $t(218) = 3.31$, $p = .001$, $d = 0.22$.

On the first-recall test, participants reported the harder response numbers more often. Figure 5B shows the frequency of participants' reports of the "first number of response circles to come to mind" for the easy-risky and hard-risky doors. For the easy-risky door, significantly more participants reported the harder number (3) than the easier number (1), χ^2 (1, N = 199) = 4.83, $p = .028$. For the hard-risky door, significantly more participants reported the hardest number (9) than the non-extreme number (7), χ^2 (1, N = 198) = 40.9, $p < .001$.

Responses on this memory test correlated significantly with choice behavior for both risky doors. Figure 5C plots risky choices on the risk-preference trials separated by responses on the first-recall test. For the easy-effort option, people who reported 1 response showed a higher percentage of risky choices (66.6 ± 3.3%; N=84) than those who reported 3 responses (31.8 ± 2.9%; N = 115), $t(197) = 7.88$, $p < .001$, $d = 1.13$. Similarly, for the hard-effort option, people who reported 7 responses showed a higher percentage of risky choices (58.8 ± 4.1%; N = 54) than those who reported 9 responses (29.1 ± 2.7%; N = 144), $t(196) = 6.65$, $p < .001$, $d = 1.06$. Partial correlations between first outcomes reported and risky choice were significant, when controlling for the obtained average outcomes of the risky options (easy: $r_p\ (196) = -.46$, $p < .001$; hard: $r_p\ (195) = -.39$, $p < .001$).

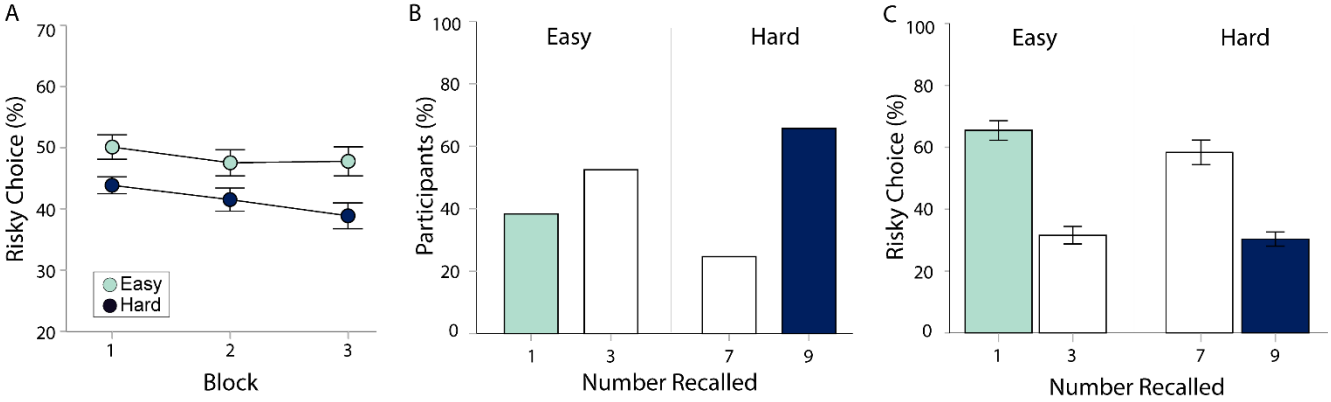

Figure 5. Results of Experiment 2. A. Percentage (± SEM) of risky choices for the decisions involving easy or hard options for each block of choice trials. B. Percentage of participants who responded with 1 or 3 for the easy-risky door, and with 7 or 9 for the hard-risky door on the first-recall test. C. Mean percentage of risky choices (±SEM) for the decisions involving easy or hard options, split by answer on the first-recall test. In both

panels B and C, green bars indicate the low extreme, navy bars indicate the high extreme, and white bars indicate non-extreme values.

Experiment 3

In both of the first two experiments, the number of responses participants made and the time taken to complete the responses both varied with effort level. This covariation simulates many real-world situations in which time and effort are correlated (walking the long route is more effortful and takes longer; solving a hard math problem to completion usually takes more time than solving an easy problem). Increases in effort, however, do not always require an increase in time. One can work out on a treadmill for a fixed amount of time at a high pace or a low pace. A cashier may spend their working hours serving many or few customers. Experiment 3 was designed to assess whether the results from the first two experiments would replicate if time was controlled so that it did not vary substantially across effort levels.

Methods

Participants

We recruited 139 participants from Prolific Academic. Participants were paid £7 for completing the experiment. They were informed that they needed to complete 108 choice trials plus some memory and preference tests in order to earn their completion code and that the task should take approximately 45 min to complete. Three participants were excluded because they were either not recorded on Prolific (N = 1) or exceeded the time limit of 115 min (N = 2). These exclusions left 136 participants (76 males, 60 females, age range of 18 to 62 with a mean of 35.5 [SD = 11.5]).

Procedure

The procedure was the same as that used in Experiment 2 with four exceptions. First, the required number of clicks following the choice doors was the same as in Experiment 1 (see Table 1). Second, the delay prior to each sequentially presented response circle was reduced to 0.1 s. Third, a delay was inserted following the response to the last sequentially presented circle in order to equate average trial duration across effort levels. To make the trial duration less predictable, this delay was adjusted so that the total trial duration had a mean of 10 s and a range of 8 to 12 s (in increments of 0.25 s). This duration spanned from the onset of the first response circle to the presentation of an X, centered on the screen, that needed to be clicked to start the next trial. If participants failed to complete all of the responses in the scheduled time, they were still allowed to finish, and then a 1-s delay was presented after the last click before the X appeared to indicate the next trial. Fourth, in this experiment there was one training block with eight single-option trials followed by five blocks that each provided four single-option trials (one for each door), eight catch trials (four for each type of choice), and eight decision trials (four for each type of choice) for a total of 20 trials per block.

After the choice trials, all participants were given memory-recall and frequency-estimation tests similar to those described in Experiment 1.

Results

We again used choices on effort-preference trials to partition participants into low-effort choosers (chose easy doors on 60% or more of the effort-preference trials) and high-effort choosers (chose hard doors on 60% or more of the effort-preference trials), resulting in 103 low-effort choosers and seven high-effort choosers. Results for the high-effort choosers are presented in the Supplemental Materials. The results reported below are for the 103 participants who chose

easy options on effort-preference trials. All *t*-tests were pre-registered and are reported as one-sided.

On risk-preference trials, people were again more risk averse for hard options than for easy options, even with the trial duration fixed. Figure 6 shows the percentage of risky choices made when participants chose between an easy-fixed door and an easy-risky door, or between a hard-fixed door and a hard-risky door across blocks of choices. Averaged over the last two blocks, participants chose the risky option 11.0 ± 4.2 percentage points less often for the hard-effort decision (42.1 ± 3.1%) than for the easy-effort decision (53.2 ± 3.7%), $t(102) = 2.63$, $p < .01$, $d = 0.26$.

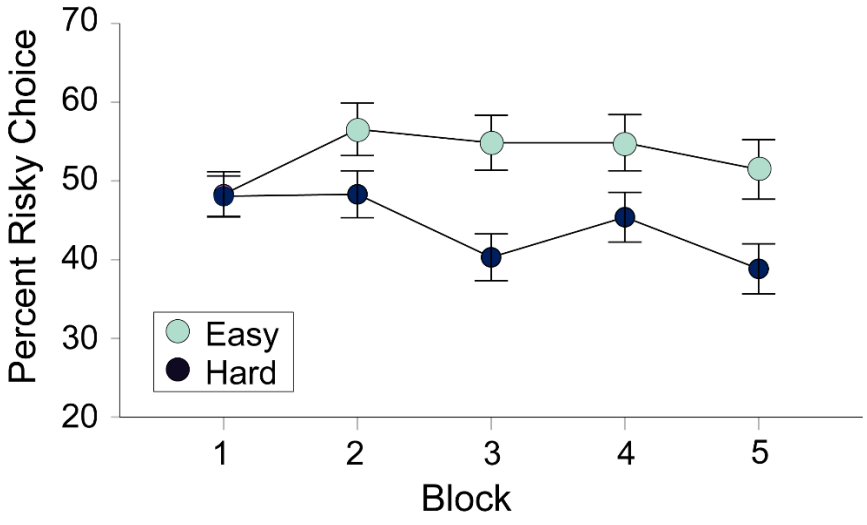

Figure 6. Risky choice results for Experiment 3. Mean percentage (± SEM) of risky choices for the decisions involving easy or hard options for each block of choice trials.

Figure 7A shows the frequency of participants' reports of the "first number of response circles to come to mind" for the easy-risky and hard-risky doors. For the easy-risky door, significantly more participants reported 1 than 5, χ^2 (1, N = 94) = 8.34, $p = .004$. For the hard-risky door, significantly more participants reported 11 than 7, χ^2 (1, N = 89) = 39.1, $p < .001$.

Thus, participants were more likely to report the numbers at ends of the distribution (extreme easy or extreme hard) as the first number to come to mind for the risky doors.

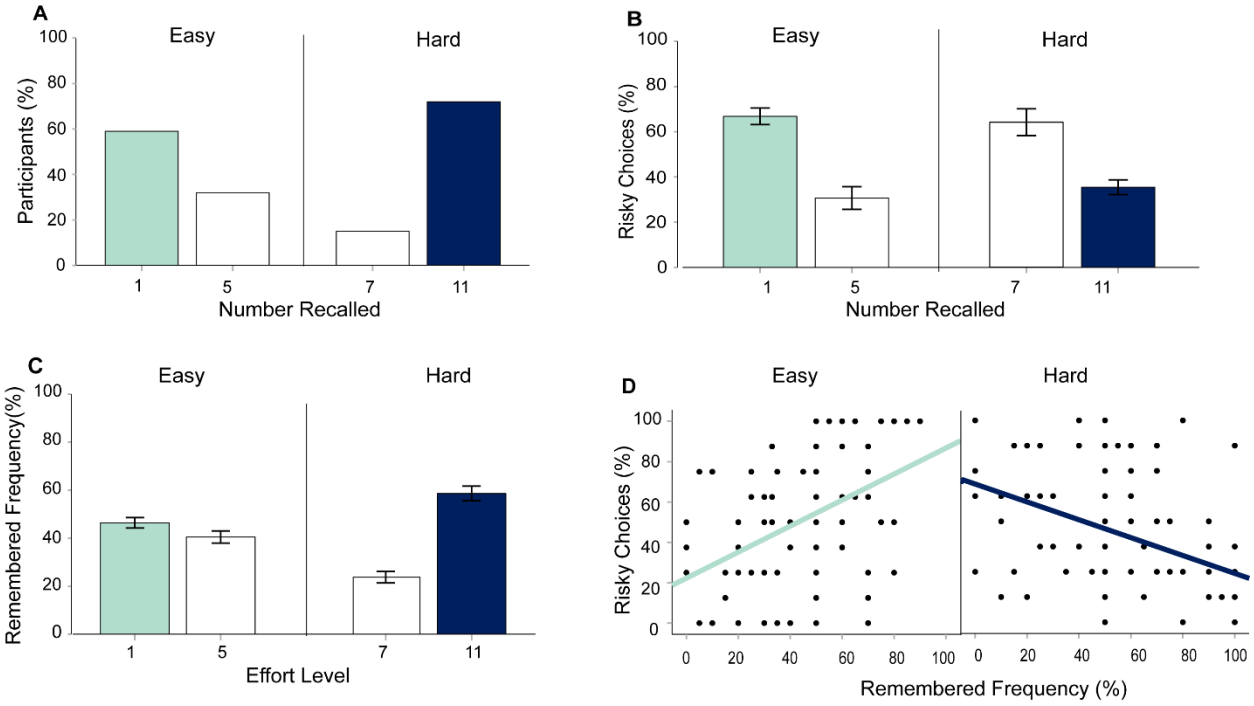

Figure 7. Results of the memory tests and correlations with risky choice in Experiment 3. (A) Percentage of participants who responded with 1 or 5 for the easy-risky door, and with 7 or 11 for the hard-risky door on the first-recall test. (B) Mean risk preference (±SEM) for easy-effort and hard-effort choices, split by answer on the first-recall test. (C) Mean percentage (±SEM) reported on the remembered frequency test that 1 or 5 response circles occurred on the easy-risky door and that 7 or 11 response circles occurred on the hard-risky door. (D) Scatterplot of risk preference on easy-effort decisions as a function of remembered frequency of the easiest outcome (1 response) and risk preference on hard-effort decisions as a function of remembered frequency of the hardest outcome (11

responses). Each dot represents an individual participant, and the lines indicate the linear regression.

Figure 7B plots risky choices on the risk-preference trials separated by responses on the first-recall test. For the easy-effort option, people who reported 1 response showed a higher percentage of risky choices (66.8 ± 3.7%; N=61) than those who reported 5 responses (30.7 ± 5.0%; N = 33), $t(92) = 5.81$, $p < .001$, $d = 1.26$. Similarly, for the hard-effort option, people who reported 7 responses showed a higher percentage of risky choices (64.2 ± 6.0%; N = 15) than those who reported 11 responses (35.5 ± 3.2%; N = 74), $t(87) = 3.75$, $p < .001$, $d = 1.06$. The partial correlations between first outcomes reported and risky choice were significant, when controlling for the obtained average outcomes of the risky options (easy: $r_p\,(91) = -.42$, $p < .001$; hard: $r_p\,(86) = -.29$, $p < .01$).

Figure 7C shows the mean reported frequency (in percent of trials) of 1 or 5 responses for the easy-risky door and of 7 or 11 responses for the hard-risky door. For the easy-risky door, participants reported a slightly higher frequency (5.93 ± 4.2%) of occurrence for the extreme (1) number of responses than for the non-extreme (5) number of responses, but this result was not statistically significant, $t(95) = 1.40$, $p = .17$, $d = .14$. For the hard-risky door, participants reported the extreme number (11) of responses as having occurred 34.9 ± 4.9 percentage points more often than the non-extreme number (7) of responses, $t(94) = 7.06$, $p < .001$, $d = .73$.

Figure 7D plots risk preference in the last 2 blocks against remembered frequency of the extremes (1 or 11 responses). For the easy-effort decisions, risky choices increased with judged frequency of the easy extreme (1 response), $r\,(94) = .41$, $p < .001$, even when controlling for outcomes experienced, $r_p\,(93) = .34$, $p < .001$. For the hard-effort decisions, risky choices decreased with judged frequency of the hardest extreme (11 responses), $r\,(93) = -.45$, $p < .001$,

even when controlling for outcomes experienced, r_p (92) = -.41, p <.001. Thus, individual differences in the remembered frequency of the different amounts of effort correlated with risky choice for decisions involving both easy and hard options.

Discussion

These experiments add a new dimension of effort risk into the examination of effort-based decision-making. The studies explored the basic question of how people choose between options that lead to the same reward but differ in the effort required and the riskiness of this effort. Previously research on effort-based choice has focused primarily on how effort discounts rewards (Botvinick et al., 2009; Hartmann, 2013) and trades off with reward (e.g., Klein-Flügge, et al., 2016; Treadway, Bossaller, et al., 2012); however, there are many situations where the outcome of a choice is constant but the effort required to obtain it is uncertain.

The set of three studies also addressed whether experience-based choice for risky effort would show biases in risk preference and memory similar to those that have been found for experience-based choice for risky reward (e.g., Ludvig et al., 2014; Madan et al., 2014). People showed clear biases in both risk preference and their memory for effort. In all three experiments, people were more risk averse for decisions involving hard-effort (worse) outcomes than for decisions involving easy-effort (better) outcomes, paralleling findings with risky reward. This result held both when time to complete each trial varied with the effort level (Experiments 1 and 2), and when time was controlled so that it was similar across effort levels (Experiment 3). Similar to results with experience-based risky choice for rewards, peoples' risky choice showed considerable variation between individuals, but this individual variation was strongly correlated with their responses on the memory tests. Large individual differences have also been found on other effort-based tasks (Treadway, Buckholtz et al., 2012).

For risky rewards, memory tests have found that people overweight the extreme outcomes (best and the worst rewards). Specifically, people are more likely to report the extremes of the experienced range as the first outcome to come to mind on recall tests, and they overestimate the frequency of extreme outcomes (best and worst) relative to equally-often experienced non-extreme outcomes (Madan et al., 2014; 2017). These effects in memory for reward are typically strongest and most consistent for the worst outcomes (i.e., relative losses; see Madan et al., 2019). For risky effort, it appears that people are also most likely to overweight the worst outcome, but in this case the worst outcome is the one requiring the most effort (highest number of clicks). On memory tests across experiments, people were more likely to recall, and they over-estimated the frequency of, the hardest outcome. Results for the memory tests were not consistent across experiments for easy outcomes. Thus, while prior work on memory for rewards suggests overweighting of both extremes with more overweighting of the worst extreme, the current studies on memory for effort provides consistent evidence only for overweighting of the hardest work.

A tendency to overestimate the hardest effort may be consistent with research on "overclaiming" (Schroeder, Caruso & Epley, 2016) in which group members' estimations of their contributions to team work sums to greater than 100%; this overestimation suggests that people remember working harder than other members of a team, referred to as an egocentric bias (e.g., Ross & Sicoly, 1979), perhaps due to an availability heuristic (Tversky & Kahneman, 1973) in which one's own hard work is more readily recalled than the effort made by others. The finding that people showed memory biases for the highest effort they exerted may also have implications for industrial psychology. If people are more likely to remember the times they had to work hard than the times they had it easier, this bias could impact not only job satisfaction, but

also how willing people are to risk the possibility of having to work harder to find potentially better ways to achieve an outcome. In cases where potentially more effortful innovation is desirable, it might be necessary to provide facilitative measures, such as reminder cues of the better possible outcome of a risky choice (Ludvig, Madan & Spetch, 2015).

These studies show that memory for the outcomes of the risky option was a reliable correlate of individual differences in risk sensitivity. Those who recalled the harder response number, and those who judged the harder response number as having occurred more often, were less likely to choose the risky option. In other words, people who remembered the harder work avoided options that could potentially lead to the harder work. Although the evidence for this relationship is correlational, and therefore causality cannot be inferred, these results provide strong evidence for the inter-relation between risky choice and effort memory, consistent with findings from risky choice for amount (Madan et al., 2014; 2017). These results suggest that models of choice should consider the relationship between memory and choice for risky decisions involving effort as well as reward. In the case of effort, the hardest work seems to particularly stand out.

References

Apps, M. A., Grima, L. L., Manohar, S., & Husain, M. (2015). The role of cognitive effort in subjective reward devaluation and risky decision-making. *Scientific Reports*, *5*(1), 1-11.

Berger-Tal, O., Mukherjee, S., Kotler, B. P., & Brown, J. S. (2009). Look before you leap: Is risk of injury a foraging cost? *Behavioral Ecology and Sociobiology*, *63*(12), 1821-1827.

Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 16-27.

Charnov, E. L. (1976). Optimal Foraging, the Marginal Value Theorem. *Theoretical Population Biology*, *9*, 129-136.

Eddy, P., Heckenberg, R., Wertheim, E. H., Kent, S., & Wright, B. J. (2016). A systematic review and meta-analysis of the effort-reward imbalance model of workplace stress with indicators of immune function. *Journal of Psychosomatic Research*, *91*, 1-8.

Fantino, E. (1967). Preference for mixed- versus fixed-ratio schedules. *Journal of the Experimental Analysis of Behavior, 10*, 35-43.

Hernandez Lallement, J., Kuss, K., Trautner, P., Weber, B., Falk, A., & Fliessbach, K. (2013). Effort increases sensitivity to reward and loss magnitude in the human brain. *Social Cognitive and Affective Neuroscience*, *9*(3), 342-349.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517-523.

Gold, J. M., Strauss, G. P., Waltz, J. A., Robinson, B. M., Brown, J. K., & Frank, M. J. (2013). Negative symptoms of schizophrenia are associated with abnormal effort-cost computations. *Biological Psychiatry*, *74*(2), 130-136.

Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. *Behavioural Processes*, *100*, 192-196.

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Sciences*, *22*(4), 337-349.

Jensen, G. D. (1963). Preference for bar pressing over" freeloading" as a function of number of rewarded presses. *Journal of Experimental Psychology*, *65*(5), 451-454.

Kacelnik, A., & Marsh, B. (2002). Cost can increase preference in starlings. *Animal Behavior, 63*, 245-250.

Kacelnik, A., & Bateson, M. (1996). Risky theories—the effects of variance on foraging decisions. *American Zoologist*, *36*, 402-434.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–292.

Kirshenbaum, A. P., Szalda-Petree, A. D., & Haddad, N. F. (2000). Risk-sensitive foraging in rats: the effects of response-effort and reward-amount manipulations on choice behavior. *Behavioural Processes*, *50*(1), 9-17.

Kivetz, R. (2003). The effects of effort and intrinsic motivation on risky choice. *Marketing Science*, *22*(4), 477-502.

Klein-Flügge, M. C., Kennerley, S. W., Friston, K., & Bestmann, S. (2016). Neural signatures of value comparison in human cingulate cortex during decisions requiring an effort-reward trade-off. *Journal of Neuroscience*, *36*(39), 10002-10015.

Konstantinidis, E., Taylor, R. T., & Newell, B. R. (2018). Magnitude and incentives: Revisiting the overweighting of extreme events in risky decisions from experience. *Psychonomic Bulletin & Review*, *25*(5), 1925–1933.

637 Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the
638 avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139*(4),
639 665-682.
640 Krebs, J. R. (1980). Optimal foraging, predation risk and territory defence. *Ardea*, *68*, 83-90.
641 Lagorio, C. H., & Hackenberg, T. D. (2010). Risky choice in pigeons and humans: A cross-
642 species comparison. *Journal of the Experimental Analysis of Behavior*, *93*(1), 27-44.
643 Ludvig, E. A., & Spetch, M. L. (2011). Of black swans and tossed coins: Is the description-
644 experience gap in risky choice limited to rare events? *PLOS ONE, 6,* e20262.
645 Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2014). Extreme outcomes sway risky decisions
646 from experience. *Journal of Behavioral Decision Making. 27*, 146–156.
647 Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2015). Priming memories of past wins induces risk
648 seeking. *Journal of Experimental Psychology: General*, *144*(1), 24-29.
649 Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2014). Remembering the best and worst of times:
650 Memories for extreme outcomes bias risky decisions. *Psychonomic Bulletin &*
651 *Review*, *21*, 629-636.
652 Madan, C.R., Ludvig, E.A., & Spetch, M. L. (2017). The role of memory in distinguishing risky
653 decisions from experience and description. *Quarterly Journal of Experimental*
654 *Psychology, 70*, 2048–2059.
655 Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2019). Comparative inspiration: From puzzles with
656 pigeons to novel discoveries with humans in risky choice. *Behavioural Processes*, *160*,
657 10-19.
658 Madan C. R., Spetch, M. L, Machado, F.M.D.S, Mason, A. & Ludvig, E. A. (2021). Encoding
659 context determines risky choice. *Psychological Science, 32,* 743-754.

Meyer, S. F., Schley, D. R., & Fantino, E. (2011). The role of context in risky choice. *Behavioural Processes*, *87*(1), 100-105.

Navarro, J., & Osiurak, F. (2015). When do we use automatic tools rather than doing a task manually? Influence of automatic tool speed. *The American Journal of Psychology*, *128*(1), 77-88.

Nagengast, A. J., Braun, D. A., & Wolpert, D. M. (2011). Risk-sensitivity and the mean-variance trade-off: decision making in sensorimotor control. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1716), 2325-2332.

Norton, M. I., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology*, *22*(3), 453-460.

Osborne, S. R. (1977). The free food (contrafreeloading) phenomenon: A review and analysis. *Animal Learning & Behavior*, *5*(3), 221-235.

Otto, A. R., & Daw, N. D. (2019). The opportunity cost of time modulates cognitive effort. *Neuropsychologia*, *123*, 92-105.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behaviour Research, Methods 51*, 195–203.

Platt, M. L., & Huettel, S. A. (2008). Risky business: The neuroeconomics of decision making under uncertainty. *Nature Neuroscience*, *11*(4), 398-403.

Reyna, V. F., & Lloyd, F. J. (2006). Physician decision making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology: Applied*, *12*(3), 179-195.

Rosenberger, K., Simmler, M., Nawroth, C., Langbein, J., & Keil, N. (2020). Goats work for food in a contrafreeloading task. *Scientific Reports*, *10*(1), 22336.

Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, *37*(3), 322-336.

Simianu, V. V., Grounds, M. A., Joslyn, S. L., LeClerc, J. E., Ehlers, A. P., Agrawal, N., ... & Flum, D. R. (2016). Understanding clinical and non-clinical decisions under uncertainty: A scenario-based survey. *BMC Medical Informatics and Decision Making*, *16*(1), 1-9.

Salamone, J. D., Correa, M., Yang, J-H., Rotolo, R., & Presby, R. (2018). Dopamine, effort-based choice, and behavioral economics: Basic and translational research. *Frontiers in Behavioral Neuroscience, 12*, 52.

Schroeder, J., Caruso, E. M., & Epley, N. (2016). Many hands make overlooked work: Over-claiming of responsibility increases with group size. *Journal of Experimental Psychology: Applied*, *22*(2), 238-246.

Sullivan-Toole, H., DePasque, S., Holt-Gosselin, B., & Galván, A. (2019). Worth working for: The influence of effort costs on teens' choices during a novel decision making game. *Developmental Cognitive Neuroscience*, 100652.

Stokes, L. J. G., Davies, A., Lattimore, P., Winstanley, C., & Rogers, R. D. (2018). Exploring preferences for variable delays over fixed delays to high-value food rewards as a model of food-seeking behaviours in humans. *Philosophical Transactions of the Royal Society B*, *374*(1766), 20180141.

Tarte, R. D. (1981). Contrafreeloading in humans. *Psychological Reports*, *49*(3), 859-866.

Treadway, M. T., Buckholtz, J. W., Cowan, R. L., Woodward, N. D., Li, R., Ansari, M. S., ... & Zald, D. H. (2012). Dopaminergic mechanisms of individual differences in human effort-based decision-making. *Journal of Neuroscience*, *32*(18), 6170-6176.

Treadway, M. T., Bossaller, N. A., Shelton, R. C., & Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: A translational model of motivational anhedonia. *Journal of Abnormal Psychology*, *121*(3), 553-558.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207-232.

Vis, B., & Van Kersbergen, K. (2007). Why and how do political actors pursue risky reforms? *Journal of Theoretical Politics*, *19*(2), 153-172.

Zentall, T.R. (2010) Justification of effort by humans and pigeons: Cognitive dissonance or contrast? *Current Directions in Psychological Science. 19*, 296–300.

Acknowledgements

This research was funded by the Alberta Gambling Research Institute (AGRI) and a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant to Marcia L. Spetch. Alice Mason was supported by a Leverhulme Early Career Fellowship (ECF-2018-408). Door images were extracted from "Irish Doors" on fineartamerica.com with permission from Joe Bonita. We thank Yang Liu and Sucheta Chakravarty for assistance with programming, Maria Shanks for assistance with testing participants, and Sarah Treit at FiguresFirst.ca for assistance with data figures. Correspondence regarding this article should be addressed to Alice Mason at Alice.M.A.Mason@warwick.ac.uk.

Supplemental Materials

Experiment 2: Circle-Preference and Pattern-Association Tests.

At the end of the experiment, participants were given two tests to assess their preference and learning about the patterns. First, a *Circle-Preference* test assessed whether participants preferred patterns associated with low effort. The six circle patterns were shown simultaneously in randomized screen locations and participants were instructed to *'Click on your MOST preferred circle.'* After a circle was clicked, the screen went blank for 1 s and then the six circle patterns were again shown simultaneously in newly randomized screen locations, along with the instruction: *'Click on your LEAST preferred circle.'* The second test was a check that participants learned the associations between the patterns and the associated effort levels. On the *Pattern-Association* test, each circle pattern was presented one at a time (in random order for each participant) and the participant was instructed: *"Type the number of times you had to click this circle each time it came up".*

The 219 participants who chose the low-effort options on effort-preference trials showed a strong preference for circle patterns associated with less effort on the Circle-Preference test. For each participant, we calculated a single circle-preference score based on the difference between the number of responses associated with the most preferred circle pattern and the number of responses associated with their least preferred circle pattern. Positive difference scores indicate preference for circles associated with higher effort whereas negative difference scores indicate preference for circles associated with lower effort. The mean difference score was significantly below 0 (-3.7 ± 0.3%), $t(218) = 11.7$, $p < .001$, $d = 0.79$, indicating strong preference for stimuli associated with lower effort. The pattern-association test confirmed that these participants learned the associations between circle patterns and effort levels. Participants

showed a strong linear relationship between number reported and the associated effort level of the circle pattern. Finally, we tested the consistency of effort preferences by conducting a correlation between choice of the hard options on effort-preference trials and the circle-preference score using all 237 participants (i.e., including those who preferred higher effort). There was a significant partial correlation between these values, $r_p(234) = .23$, $p < .001$, even when controlling for differences in exposure to the most and least preferred circle patterns.

Results for High-Effort Choosers

High-effort choosers were defined as participants who chose the hard options in 60% or more of the effort-preference trials. There were 15 high-effort chooser in Experiment 1, 6 in Experiment 2 and 7 in Experiment 3. In each experiment, the high-effort choosers showed the opposite pattern of risk preference to the low-effort choosers, choosing the risky option more often for hard options (49.6%, 59.2%, and 53.5%) than for easy options (40.0%, 51.3%, and 41.1%) for Experiments 1 to 3 respectively. These differences were not significant in any of the experiments, but the sample sizes were very small.

On the recall test for the easy-risky door, the extreme low-effort number was recalled by 2, 2, and 5 participants and the non-extreme low-effort number was reported by 1, 2 and 0 participants in the three experiments. For the hard-risky door, the extreme number was reported by 1, 2, and 4 participants, and the non-extreme number was reported by 1, 0, and 1 participants across the three experiments.

The pattern preference and recall tests of Experiment 2 showed that high-effort choosers preferred the patterns associated with high effort, and their difference score (6.0±1.0 %; Mean ±SEM) was significantly above zero, $t(5) = 6.0$, $p = .002$, $d = 2.45$. This result was opposite to the preference shown by the low-effort choosers (see below). The high-effort choosers were

772 generally accurate in reporting the effort level associated with each circle pattern, showing a
773 significant linear trend in reported number for patterns associated with 1, 2, 3, 7, 8, and 9
774 required responses, respectively, $F(1,5) = 10.8$, $p = .022$, $\eta_p^2 = .68$.
775
776 **Analyses of risky choice by blocks for all experiments**
777 In all experiments, risky choice was based on the last block(s) of choice trials, so as to
778 measure preferences once contingencies have been learned. We pre-registered t-tests to compare
779 risky choice on easy and hard decisions in those final blocks(s). Here, we present additional
780 exploratory analyses of variance (with Greenhouse-Geisser correction) on risky choice by block
781 of choice trials for the three experiments.
782 In Experiment 1, there was a significant main effect of block, $F(3,178)=6.25$, $p<.001$, η_p^2
783 $=.10$. People chose the risky option more often for the easy options compared to the hard options
784 but the main effect of effort was marginally not significant, $F(1,59)=3.65$, $p=.061$, $\eta_p^2 = .058$. For
785 the hard options people chose the risky option less across the experiment but for the easy options
786 they chose the risky option more as the experiment progressed (significant interaction between
787 block and effort level, $F(4,185) =4.07$, $p<.01$, $\eta_p^2 = .065$).
788 In Experiment 2, people tended to select the risky option less often across the experiment
789 (main effect of block, $F(2,385)=3.58$, $p=.03$, $\eta_p^2 =.016$), and they were more risk seeking for
790 easy options (main of effort, $F(1,218)=10.81$, $p = .001$, $\eta_p^2 = .047$), but there was no significant
791 interaction between block and effort level, $F(2,407) =0.88$, $p=.41$, $\eta_p^2 = .004$.
792 In Experiment 3, people were more risk seeking for the easy options compared to the
793 hard options across the blocks (main effect of effort, $F(3,178)=7.46$, $p<.01$, $\eta_p^2 =.068$), but their
794 risk preferences did not change significantly across the blocks (no significant main effect of

795 block, $F(4,350.35)=2.16$, p=.083, $\eta_p^2 = .021$). There was a significant interaction between block
796 and effort level, $F(4,387) =2.51$, p=.04, $\eta_p^2 =.024$. For the hard options people chose the risky
797 option less as the experiment progressed, whereas for the easy option people initially chose the
798 risky option more often.
799
800 **Supplemental Experiment**
801 Between conducting Experiments 2 and 3, we pre-registered and ran the following
802 version of the time-controlled experiment. Because more participants than we anticipated failed
803 to complete the trials within the scheduled limit, thereby defeating the time-controlled aspect of
804 the experiment, we then altered the details of the task and ran another time-controlled
805 experiment, reported as Experiment 3.
806 **Methods**
807 ***Participants***
808 We recruited 160 participants from Prolific Academic (74 males, 84 females, age range
809 of 19 to 65 with a mean of 32.7). One participant did not report their age, and two participants
810 did not report a gender. Participants were paid £7 for completing the experiment. They were
811 informed that they needed to complete 108 choice trials plus some memory and preference tests
812 in order to earn their completion code and that the task should take approximately 45 min to
813 complete.
814 ***Procedure***
815 The procedure was identical to that used in Experiment 3 with three exceptions. First, the
816 required number of clicks following the choice doors differed. The number of circles participants
817 needed to click was 4 for the easy fixed door, 2 or 6 for the easy risky door, 14 for the hard fixed

door and 12 or 16 for the hard risky door. Second, the variable trial time limit was set to 10-15 s (in increments of .25 s). Third, there was no effort-liking scale at the end.

Results

There were two exclusion criteria: 1) fewer than 60% choice of the easy-effort option on catch trials and 2) more than 10 trials on which the trial was not completed within the programmed time limit. 57 participants were thus excluded, leaving 103 participants.

Figure S1 shows the percentage of risky choices made when participants chose between an easy-fixed door and an easy-risky door, or between a hard-fixed door and a hard-risky door across blocks of choices. As per the pre-registration, averaged over the last two blocks, participants were mildly more risk-seeking for easy-effort decisions (47.7 ± 3.50%) than for the hard-effort decisions (43.5 ± 3.05%), but this difference was not significant, $t(102) = 0.90$, $p=.90$, $d = 0.09$. People chose, however, reliably more riskily on easy-effort trials across the whole experiment, however, as confirmed by an exploratory ANOVA on risky choice by blocks [Main effect of effort: $F(1, 102) = 4.08$, $p=.046$, $\eta_p^2 = .038$]. There was no significant change across blocks, $F(3.33,340.15)=0.675$, $p=.583$, $\eta_p^2=.007$, and no significant interaction between effort and block, $F(3.53,360.58)=1.66$, $p=.167$, $\eta_p^2 = .016$.

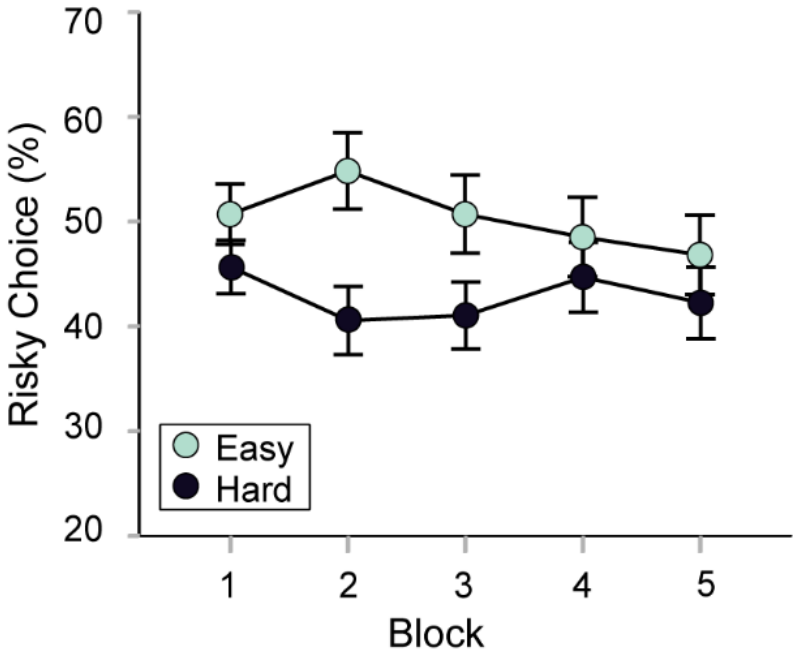

Figure S1. Risky choice results for Experiment S1. Mean percentage (± SEM) of risky choices for the decisions involving easy or hard options for each block of choice trials.

Figure S2A shows the percentage of participants who reported 2 or 6 as the "first number of response circles to come to mind" for the easy-risky door and 12 or 16 for the hard-risky door. For the easy-risky door, more participants reported 2 (the low extreme) than 6, but the difference was not significant, χ^2 (1, N = 91) = 3.18, p = .075. For the hard-risky door, significantly more participants reported 16 (the high extreme) than 12, χ^2 (1, N = 84) = 19.1, p < .001.

Figure S2B plots risky choices on the risk-preference trials separated by responses on the first-recall test. For the easy-effort decision, people who reported 2 responses showed a higher percentage of risky choices (55.3 ± 4.9%; N=54) than those who reported 6 responses (36.8 ± 5.0%; N = 37), t(89) = 2.46, p =.016, d =.53. Similarly, for the hard-effort decision, people who reported 12 responses showed a higher percentage of risky choices (61.9 ± 6.2%; N = 22) than those who reported 16 responses (37.7 ± 3.9%; N = 62), t(82) = 3.21, p < .002, d =.80.

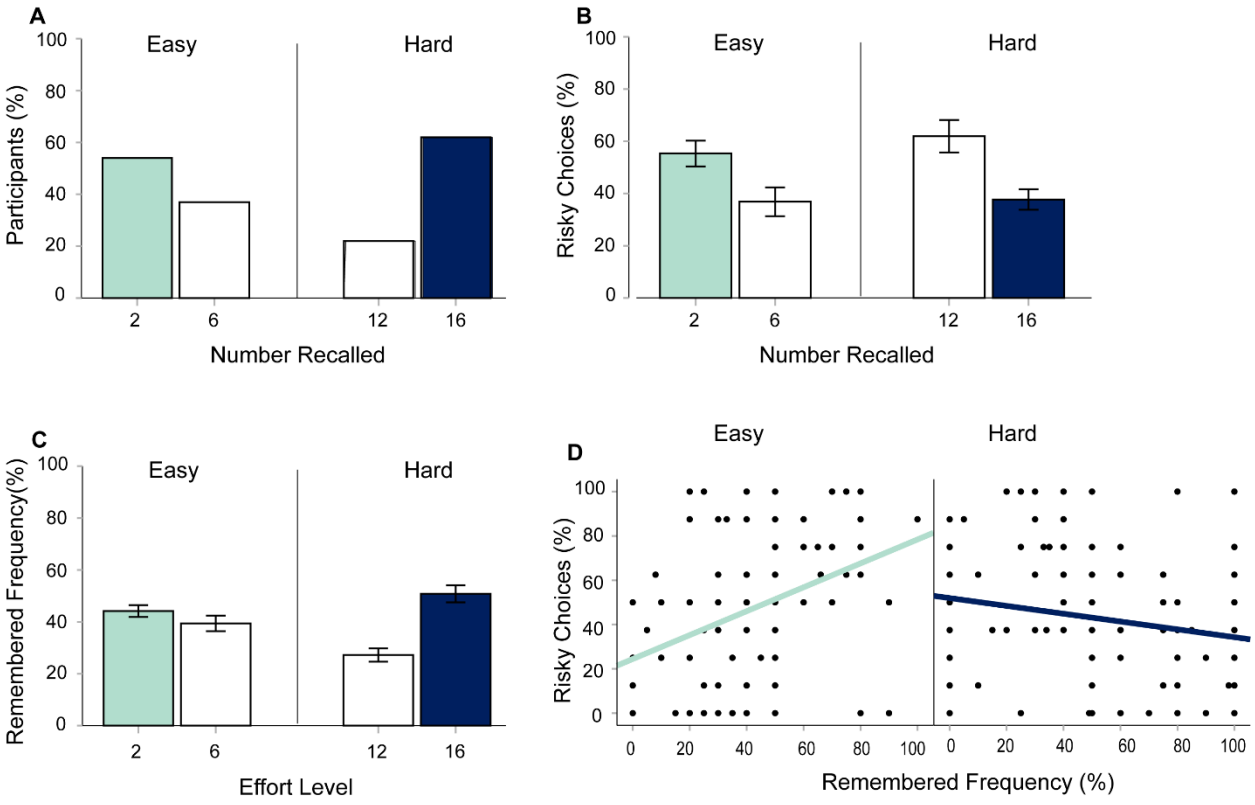

Figure S2. Results of the memory tests and correlations with risky choice in Experiment S1. (A) Percentage of participants who responded with 2 or 6 for the easy-risky door, and with 12 or 16 for the hard-risky door on the first-recall test. (B) Mean risk preference (±SEM) for easy-effort and hard-effort choices, split by answer on the first-recall test. (C) Mean percentage (±SEM) reported on the remembered-frequency test that 2 or 6 response circles occurred on the easy-risky door and that 12 or 16 response circles occurred on the hard-risky door. (D) Scatterplot of risk preference on easy-effort decisions as a function of

remembered frequency of the easiest outcome (2 responses) and risk preference on hard-effort decisions as a function of remembered frequency of the hardest outcome (16 responses). Each dot represents an individual participant, and the lines indicate the linear regression.

Figure S2C shows the mean reported frequency (in percent of trials) of 2 or 6 responses for the easy-risky door and of 12 or 16 responses for the hard-risky door. Only participants who provided frequency estimates were included in the analysis. For the easy-risky door, participants did not report a significantly higher frequency (4.77 ± 4.75%) of occurrence for the extreme (2) number of responses than for the non-extreme (6) number of responses, $t(93) = 1.00$, $p = .32$ $d = 0.10$. For the hard-risky door, participants reported the extreme number (16) of responses as having occurred 24.6 ± 5.2 percentage points more often than the non-extreme number (12) of responses, $t(95) = 4.54$, $p < .001$, $d = 0.46$.

Figure S2D plots risk preference in the last 2 blocks against remembered frequency of the extremes (2 or 16 responses). For the easy-effort decisions, risky choices increased significantly with judged frequency of the easy extreme (2 responses), $r\ (92) = .35$, $p < .001$, even when controlling for outcomes experienced, $r_p\ (91) = .32$, $p = .002$. For the hard-effort decisions, risky choices decreased with judged frequency of the hardest extreme (16 responses), but neither the correlation nor the partial correlation were significant, $r\ (94) = -.18$, $p = .07$, $r_p\ (93) = -.18$ $p = .08$.