Exploiting Risk–Reward Structures in Decision Making Under Uncertainty

Christina Leuker, Thorsten Pachur, Ralph Hertwig, and Timothy J. Pleskac

Max Planck Institute for Human Development,

Berlin, Germany

7 November, 2017

Word count: 12471

Author Note

Christina Leuker, Thorsten Pachur, Ralph Hertwig, and Timothy J. Pleskac, Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany. Author contributions were as follows: Conceptualization: C.L., R.H., T.P., & T.J.P.; Methodology: C.L, T.P., & T.J.P.; Software: C.L.; Data collection & curation: C.L.; Formal analysis: C.L.; Writing - Original Draft C.L. & T.J.P.; Writing - Reviewing & Editing - C.L., R.H., T.P., & T.J.P.; Supervision: T.J.P. This research was supported by a fellowship to CL from the MaxNetAging Research School at the Max Planck Institute for Demographic Research in Rostock, Germany. Conflicts of interest: none. We would like to thank Jann Wäscher and Chantal Wysocki for assistance with data collection, and Susannah Goss for editing the manuscript. All gamble stimuli and data can be retrieved via the Open Science Framework (anonymous link for review). Correspondence concerning this article should be addressed to Christina Leuker, Max Planck Institute for Human Development, Center for Adaptive Rationality, Lentzeallee 94, 14195 Berlin, Germany. E-mail: leuker@mpib-berlin.mpg.de.

Abstract

People often have to make decisions under uncertainty — that is, in situations where the probabilities of obtaining a reward are unknown or at least difficult to ascertain. Because outside the laboratory payoffs and probabilities are often correlated, one solution to this problem might be to infer the probability from the magnitude of the potential reward. Here, we investigated how the mind may implement such a solution: (1) Do people learn about risk–reward relationships from the environment—and if so, how? (2) How do learned risk–reward relationships impact preferences in decision-making under uncertainty? Across three studies ($N = 352$), we found that participants learned risk–reward relationships after being exposed to choice environments with a negative, positive, or uncorrelated risk–reward relationship. They learned the associations both from gambles with explicitly stated payoffs and probabilities (Experiments 1 & 2) and from gambles about epistemic events (Experiment 3). In subsequent decisions under uncertainty, participants exploited the learned association by inferring probabilities from the magnitudes of the payoffs. This inference systematically influenced their preferences under uncertainty: Participants who learned a negative risk–reward relationship preferred the uncertain option over a smaller sure option for low payoffs, but not for high payoffs. This pattern reversed in the positive condition and disappeared in the uncorrelated condition. This adaptive change in preferences is consistent with the use of the risk–reward heuristic.

Keywords: decisions under uncertainty, adaptive cognition, risk and reward, incidental learning, ecological rationality

Exploiting Risk–Reward Structures in Decision Making Under Uncertainty

In March 2016, James Stocklas won $291 million in the Florida Powerball lottery. Most people know that winning such a huge jackpot is a pretty unlikely event. Now consider his brother, Bob Stocklas. Bob bought a ticket for the same lottery at the same time as James and won just $7 (Newsome, 2016). Most people know that while winning this kind of sum is also unlikely, it is far more likely than winning the jackpot. And, of course, most people are also painfully aware that not winning anything at all is much more likely than either of these events. While this story illustrates the strange vicissitudes of fortune, for our purposes it also illustrates just how comfortable people are with estimating the probability of winning from payoff magnitudes alone. How do people "know" how to estimate the chances of winning the lottery? Why do they specifically associate the highest payoff with the lowest probability? Here, we argue that the key to understanding how the mind generates such estimates lies in an adaptive approach to cognition. Adaptive approaches to cognition seek to understand cognition within the environmental context (Anderson, 1991; Gibson, 1979; Gigerenzer et al., 2011; Marr, 1982; Simon, 1956; Stewart et al., 2006). Regarding the lottery example, one may argue that the estimates are quite attuned to what the true relationship between payoffs and probabilities in the lottery is.

Risks and rewards, or payoffs and probabilities, are linked in many choice environments beyond the lottery. Across choice environments, probably the most frequent and recurrent link between them is an inverse relationship: The higher rewards that we desire are unlikely to be obtained (Pleskac & Hertwig, 2014). However, Pleskac and Hertwig also found that the strength of the relationship varied across different domains. Monetary gambles in casinos, for instance, showed a near perfect (though biased) inverse relationship between payoffs and probabilities. In other domains, such as where to submit a scientific manuscript (trading off impact factor against acceptance rate), the risk–reward relationship was less strong. Moreover, a risk–reward relationship is not always a given. For instance, a negative relationship between risk and reward is to be expected in economic markets only if an equilibrium has been achieved, but not in newly forming markets that have not yet reached an equilibrium.

Do these different risk–reward structures influence choice? After identifying the ecological structures in which the mind usually operates, one can try to establish how the mind comes to terms with those ecological structures (Brunswik & Kamiya, 1953; Simon, 1956). In the case of risk–reward structures, Pleskac and Hertwig (2014) found that the mind seems to exploit risk–reward structures in decisions under uncertainty — where people have to choose between options whose payoffs are known but probabilities are not (Knight, 1921; Luce & Raiffa, 1957; Wakker, 2010). In their study, the authors offered participants a gamble that gave them a chance to win \$$x$ at the cost of \$2, and asked them to estimate the probability of winning \$$x$. Different participants were asked to consider different magnitudes of x. As the magnitude of the potential payoff increased, the estimated probabilities of winning decreased — ultimately influencing what participants chose. These results imply that participants inferred the probability of winning from the magnitude of the payoff. Moreover, participants seemed to have represented the ecological relationship between risks and rewards as inverse. A direct link between the estimates and a specific representation of risk–reward structures has, however, not yet been established.

Inferring probabilities from the magnitude of the potential payoff might be an adaptive solution to decision-making under uncertainty — a solution that Pleskac and Hertwig (2014) refer to as the *risk–reward heuristic*. However, such an adaptive, ecological solution to taming uncertainty has specific requirements. Here, we investigate two of these requirements: First, the mind has to be sufficiently sensitive to fundamental relationships between the variables in an environment (Brunswik, 1955; Gigerenzer et al., 1991; Gibson, 1979; Marr, 1982; Simon, 1956; Stewart et al., 2006) or even mirror aspects of the environment (Anderson & Schooler, 1991; Shepard, 1967, 1987). How do people extract, or learn, the risk–reward structure from the environment? Second, people should be willing to harness the structure flexibly, as the ecological regularity varies across environments (Todd

& Gigerenzer, 2007). That is, they should be able to withhold from reflexively estimating a high payoff to be unlikely if appropriate (e.g., in a newly forming market). This argument can be developed further: Payoffs and (subjective) probabilities determine the value of an option, and ultimately choice. Therefore, different risk–reward environments should not only affect the estimates themselves but also decisions under uncertainty. To what extent do people adapt to different risk–reward structures when making decisions under uncertainty?

Figure 1 provides an overview of the assumed relationships between risk–reward structures and choice that we take in this paper. Next we hypothesize how learning of risk–reward structures might create a representation of the ecological structure in the mind, and how these representations can ultimately produce environment-dependent preference shifts in decisions under uncertainty. We then outline how we tested those hypotheses across three experiments, before reporting each experiment in detail.

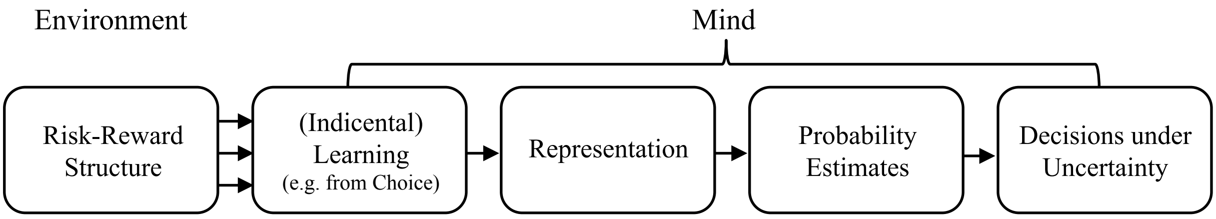

Figure 1. Summary of the assumed relationships among risk–reward structures in the world and how they ultimately shape preferences under uncertainty. All processes can be perturbed by noise.

How Can People Learn Risk–Reward Structures?

The risk–reward relationship refers to the co-occurrence of the magnitude of payoffs and probabilities. In most domains, people are not explicitly informed if a risk–reward relationship is present, nor are people explicitly informed what the relationship is. They often do not have the luxury to learn about them from explicit feedback. In this case, a risk–reward relationship would need to be acquired as people go about their primary objective when making decisions, and evaluate the options available in a choice environment. In other words, the risk–reward relationship would seem to be learned unsupervised (where there is no corrective feedback; Love, 2002), and incidentally (where learning is not the primary objective; Brooks, 1978; Dulany et al., 1984; Nelson, 1984; Ward & Scott, 1987; Wattenmaker, 1991; Whittlesea, 1987).[1]

Prior research suggests that via such incidental learning, people can be remarkably well attuned to statistical structures of their choice environments. For instance, they are quite good at estimating the frequencies of events, even when that is not their central task, and at automatically processing the frequency information (Hasher & Zacks, 1979; Hasher et al., 1987; Zacks, 2002). The ability to encode the frequency of a particular attribute is often the basis of adaptive explanations of choice patterns (e.g., Goldstein & Gigerenzer, 2002; Marewski & Schooler, 2011; Stewart et al., 2006). For example, people appear to encode the prices of goods and to use those prices later to evaluate the subjective worth of new values (Brown et al., 2008; Stewart et al., 2006; Olivola & Sagara, 2009; Ungemach et al., 2011), or use marginal distributions of either payoffs or probabilities in subjective evaluations thereof (Stewart et al., 2015; Walasek & Stewart, 2015). However, the risk–reward relationship is different from encoding and using (marginal) distributions of probabilities/frequencies and payoffs in that it requires people to learn a statistical regularity between probabilities and payoffs (i.e., a *joint* distribution). Moreover, while it is well known that people can learn associations between two variables such as a cue and a criterion (e.g., Cooksey, 1996), in preferential decision making contexts neither the probabilities nor the payoffs are the criterion.

To test people's ability to learn a risk–reward relationship in an unsupervised,

[1] One might also classify this as a case of implicit learning (see, e.g., Cleeremans et al., 1998; Frensch & Rünger, 2003; Reber, 1967, 1989; Seger, 1994; Shanks & St. John, 1994). However, a typical condition for implicit learning is that individuals lack awareness of what is learned. We are thus hesitant to use this concept, as it seems that people are aware of the risk–reward relationship (Pleskac & Hertwig, 2014).

incidental manner, we created a learning phase in which participants encountered gambles where risks and rewards that were negatively correlated, positively correlated, or uncorrelated. Across experiments, we tested participants' ability to learn from different types of gambles. Specifically, in Experiments 1 and 2, participants were asked to evaluate risky monetary gambles of the form "p chance of winning x, otherwise nothing." In Experiment 3, we examined to what extent participants learned different risk–reward structures from epistemic events when the probabilities were subjective (see also Tversky & Fox, 1995; Tversky & Wakker, 1995). Across experiments we also examined how different response types impacted learning with participants either choosing between gambles (Experiment 1) or indicating their willingness to sell for individual gambles (Experiments 2 and 3).

Finally, we also examined in what form the risk–reward relationship is represented. In Experiments 1 and 2, we asked participants if they recognized specific gambles from the earlier learning phase. In so doing, we tested whether the risk–reward structure was learned as a "risk–reward rule" or via memory of specific gambles, that is, exemplars (Erickson & Kruschke, 1998): If it was learned via exemplars, participants should be able to recognize specific gambles from the learning phase (but not similarly structured lures).

(How) Are Different Risk–Reward Structures Exploited in Decisions Under Uncertainty?

If people learn about risk–reward structures and subsequently exploit them, their decisions made under uncertainty may vary as a direct function of the risk–reward structure they have previously experienced. After the incidental learning phase, participants were asked to make a series of choices between an uncertain prospect that offered a payoff of x with an unknown probability and a sure payoff. Across the choices, we manipulated the magnitude of x. Three distinct predictions can be developed regarding how risk–reward environments should or should not impact preferences for the uncertain prospect.

If risk–reward structures are used in decisions under uncertainty to infer the values of missing probabilities from the magnitude of the payoff itself, estimated probabilities and subsequent choices regarding a given payoff magnitude should differ depending on which risk–reward structure participants previously experienced.[2] In an environment with a negative risk–reward relationship, payoffs become less and less likely as their magnitude increases. Consequently, someone using the risk–reward heuristic to infer missing probabilities will, all else being equal, avoid low (inferred) probabilities of winning high payoffs and prefer a sure outcome—provided that the sure outcome outweighs the uncertain outcome weighted by its inferred probability. Conversely, the decisions of someone who has learned that risks and rewards are positively related can be expected to show the opposite pattern. Lastly, someone who has learned that risks and rewards are uncorrelated can be expected to make decisions as if the probability estimates assigned to events were independent of their payoffs: For instance, participants may adhere to the principle of indifference and treat each outcome as equally likely and assign a probability of .5 to each outcome (Fox & Clemen, 2005; Fox & Rottenstreich, 2003).

An alternative prediction can be developed from subjective expected utility theory (Savage, 1954), according to which risk–reward structures should have no impact on people's preferences. This is because subjective expected utility theory adheres to the principle of description invariance, whereby preferences can be decomposed into the independent factors of subjective probabilities and subjective utilities. To maintain description invariance, a change in the magnitude of the outcome should not change the

[2] How does the estimation process work? In many nonlaboratory environments that have pay-to-play structure where it costs a fixed amount c to play for the possibility to earn g plus the cost of playing, the estimates can be computed as $p(g) = \frac{c}{c+g}$. In our experiments, the relationships were set to be linear, and playing incurred no costs to participants. Therefore, in an environment with a negative risk–reward relationship and n different outcomes, the probability of payoff $x_{i'}$ would be $p(x_{i'}) = 1 - x_{i'} / \max(x_i, \ldots, x_n)$. In an environment with a positive risk–reward relationship, it would be the inverse, $p(x_{i'}) = x_{i'} / \max(x_i, \ldots, x_n)$.

probability assigned to an event—otherwise, this decomposition is not possible (i.e., one's estimate of the chances of an event occurring should not be affected with it being paired with a high or low payoff; otherwise subjective probabilities can not be estimated from preferences between gambles). Thus, if subjective expected utility theory is taken as a first approximation of a descriptive theory of choice, then the larger context of the risk–reward environment should not impact people's preferences. One way description invariance could be maintained is to use the principle of indifference and treat each outcome as equally likely. Then, the principle of indifference should hold across risk–reward environments (Fox & Clemen, 2005; Fox & Rottenstreich, 2003).

Another prediction on how people deal with missing probability information in decisions under uncertainty can be derived from research on the desirability or optimism bias (Bar-Hillel & Budescu, 1995; Edwards, 1962; Irwin, 1953; Krizan & Windschitl, 2007; Sharot, 2011; Windschitl et al., 2010). According to this research, as payoffs become more desirable, they (or the event with which they are associated) are perceived as more likely. This prediction should hold irrespective of the statistical relationship between risk and reward. The affect heuristic, according to which more positive overall affect towards high payoffs can mitigate perceived risk (Pachur et al., 2012; Slovic & Peters, 2006; Slovic et al., 2004), would yield a similar prediction. That is, both the optimism bias and the affect heuristic may support the belief that—probably within limits—high payoffs are by no means unlikely.

Overview of Experiments

We conducted three experiments, each consisting of a condition-dependent learning phase and a test phase. Table 1 provides an overview of the structure of each experiment. In Experiments 1 and 2, learning environments consisted of gambles of the form "p chance of winning x, otherwise nothing." In Experiment 3, the gambles were about an epistemic event, namely, whether the maximum temperature in Berlin on a particular day in 2011 fell

Table 1

Overview of Experiments and Conditions

Experiment	Learning phase Task	Learning phase Conditions	Test phase (condition-independent)	Aim of experiment
1	Choice	Negative Uncorrelated	Decisions under uncertainty Probability estimation Recognition	Incidental learning of risk–reward structures Influence on decisions under risk and uncertainty
2	WTS	Negative Positive Uncorrelated	Decisions under uncertainty Probability estimation Recognition Payoff estimation	Incidental learning of a positive risk–reward structure Influence of response mode on learning phase task Influence on decisions under uncertainty
3	WTS	Negative risk Positive risk Negative uncertain Positive uncertain	Decisions under uncertainty Subjective probability (with payoffs) Subjective probability (without payoffs) Risk–reward task	Gambles with epistemic events Incidental learning under risk vs. uncertainty Influence on beliefs about events

Note. Learning phase stimuli were condition-dependent. All test phase tasks were condition-independent. WTS: Willingness to sell.

within a given range. This design allowed us to examine how well participants learned risk–reward structures from choice material in which the probabilities were not explicitly stated. In all three experiments, environments were constructed such that across the gambles risks and rewards were either negatively correlated, positively correlated, or uncorrelated. Importantly, participants were neither informed about the risk–reward structures nor asked to attend to them; instead they merely experienced the structure by either choosing or pricing different monetary gambles. We thus examined how well participants learned the different structures from incidental, unsupervised learning.

After the learning phase, participants completed a decisions under uncertainty task. Here, we tested how exposure to different risk–reward environments impacted participants' preferences among uncertain options. Participants then completed probability estimation tasks, which we used to test whether they had learned the risk–reward structure. In Experiments 1 and 2, they also completed a gamble recognition task that tested whether the structure was learned via memory of specific exemplars of gambles or as a rule. The order of estimation and recognition tasks was counterbalanced between participants.

Experiment 1: Do People Learn Negative vs. Uncorrelated Risk–Reward Environments and Exploit Them in Decisions Under Uncertainty?

Our first experiment had an exploratory focus. We designed it to examine how the risk–reward structure impacts decision making under both risk (probabilities given) and uncertainty (probabilities missing). To this end, we exposed participants to different risk–reward environments, asking them to make 119 choices between two nondominating gambles of the form "p chance of winning x, otherwise nothing." Between participants, the gambles were selected from one of two environments. In the negative environment, there was a negative (linear) relationship between payoffs and probabilities across all possible gambles. In the uncorrelated environment, the same payoffs and probabilities were presented, but the payoffs and probabilities were randomly paired in each gamble. We hypothesized that participants would learn about these different risk–reward structures and process the joint distribution of payoffs and probabilities as a consequence of their primary task, which was to choose the alternative they preferred.

Note that the learning phase consisted of decisions under risk (payoffs and probabilities known), and could therefore be used to test how risk–reward structure may impact decisions under risk. To do so, about halfway through this learning phase, we included gambles common to both conditions (Figure 2). Some of the common gambles we included are known to produce the 'certainty effect' (Kahneman & Tversky, 1979). However, we found very little differences between the conditions with respect to choices under risk. For instance, we found the certainty effect in both conditions. Because our article is focused on decisions under uncertainty, we do not report any further on these results. Please see the Supplemental Material for details of our hypotheses and analyses on decisions under risk.

After the learning phase, participants completed three tasks that were identical across both conditions (test phase). We hypothesized that their learning of the different risk–reward environments would shape their subsequent preferences in decisions under uncertainty. Specifically, we expected environment-dependent preferences to emerge: when choosing between an uncertain option and a sure thing, participants in the negative condition should prefer the sure thing for high payoffs and the uncertain option for low payoffs because they would infer low probabilities of winning high payoffs and high probabilities of winning low payoffs. We expected that participants in the uncorrelated condition would make payoff–independent choices (e.g., assume a probability of .5 for all payoffs). We then tested to what extent participants learned the respective risk–reward structure by explicitly asking them to estimate probabilities when presented with new payoffs. Finally, we administered a recognition task to investigate whether participants remembered specific gambles (exemplars) or whether they had extracted a "risk–reward rule" from the learning phase.

Method

Participants. Due to a lack of relevant previous studies, no data were available for a formal power calculation. We set a target sample size of 60 participants, with 30 participants per condition. In total, the sample comprised 62 adults (32 females, mean age = 25.6, SD = 3.4, proportion students = .93) from the participant pool maintained at the Max Planck Institute for Human Development (32 in the negative condition, 30 in the uncorrelated condition).[3] All reported experiments were approved by the IRB of the Max Planck Institute for Human Development. Participants gave signed informed consent prior to the experiment; they were paid a fixed rate of 10 €/hour plus a bonus contingent on their choices.

[3] Ten other participants also completed the experiment, but a coding error in the computerized experiment corrupted their data.

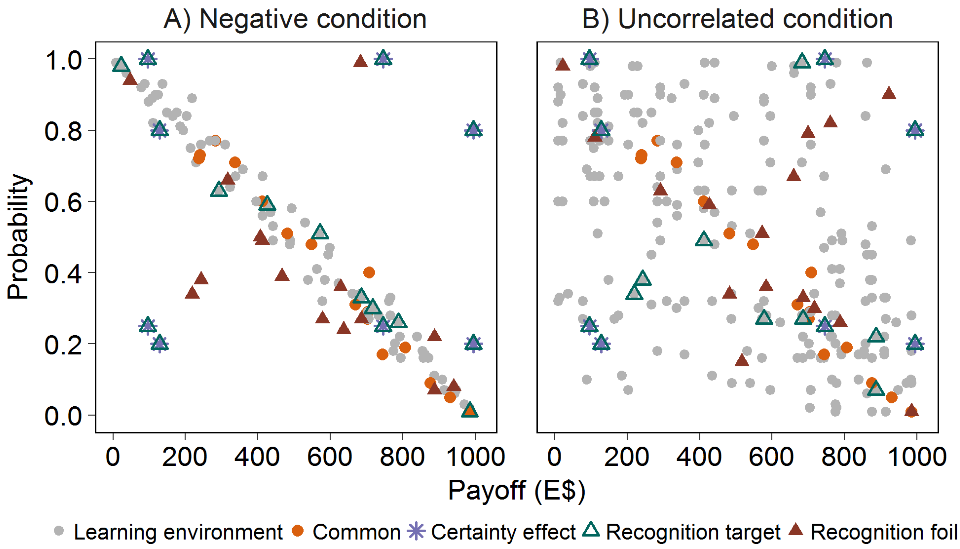

Figure 2. Stimuli used in Experiment 1. The learning phase consisted of 100 condition-dependent gamble pairs. Common gambles (10 pairs) and certainty-effect gambles (4 pairs) were randomly interspersed in the second half of the learning phase, allowing us to study condition-dependent changes in decisions under risk. Dominated options not depicted.

Decisions under risk (learning phase). During the learning phase, participants repeatedly chose between two monetary gambles of the form "p chance of winning x, otherwise nothing." All payoffs across all three were expressed using an experimental currency, E$. We did this with the goal of minimizing the impact of outside norms associated with specific currencies on the experiments.

For the negative condition, the gambles were constructed as follows. Payoffs were determined by drawing 100 random payoffs from a uniform distribution with a range 1.01–1000. The probabilities for each payoff were set so that they were inversely related to the payoff x: $p = 1 - x/1000$. We then jittered payoffs and probabilities by adding normally distributed noise with a standard deviation of 0.2 to the logit transformation of the probabilities and to the logit transformation of normalized payoffs, and we transformed those perturbed values back to the scales used in the experiment. We randomly drew 100 nondominated gamble pairs from this population of 4,950 possible problems. Differences in expected value between gambles in this condition were relatively small (Mdn_{abs} = E$67.7, $Mdn_{\%}$ = 6.8%, range E$0.17–E$281.87), as payoff always trades off with probability in a negatively correlated environment.

For the uncorrelated condition, we took the payoffs and probabilities from the 100 gamble pairs used in the negative condition, but now randomly linked probabilities and payoffs. We did this to maintain the marginal distributions of payoffs and probabilities across both the conditions and the total number of gambles in the set (see Stewart et al., 2006, 2015). We again drew 100 random gamble pairs from all possible combinations. If any of the gamble pairs had stochastically dominated options (i.e., $p_A > p_B$ and $x_A > x_B$), we switched the probabilities of gambles A and B. The expected value differences between gambles were larger in the uncorrelated condition than in the negative condition (Mdn_{abs} = E$133, $Mdn_{\%}$ = 13.3%, range E$0.52–E$844.6).

In both conditions, we included five dominated options that we used as 'catch trials' to identify participants who did not pay attention. In addition, 14 identical gamble pairs appeared in both conditions. Ten of these pairs were based on the procedure for the negatively correlated risk–reward environment. The other four pairs were designed to examine the certainty effect (see Supplemental Material). Across participants, we randomized the positions of the gambles on screen, and counterbalanced the location of payoffs and probabilities (top/bottom).

Decisions under uncertainty (test phase). To test the effects of the different risk–reward environments on subsequent decisions under uncertainty, we drew 20 random payoffs (range E$1–1000, fixed across participants) and gave them a probability of "?". Each uncertain payoff y was then matched with a half-as-large certain option (probability 100%). In a typical pair, participants thus chose between a 100% chance of winning E$50 and a "?" chance of winning E$100. We included 20 distractor gambles using another 20 random uncertain payoffs (probabilities "?"). We created smaller, certain options k by multiplying the uncertain option y with its presumed probability in a negatively correlated risk–reward environment, $k = y \times (1 - y/1000)$. The location of the uncertain and certain options on the screen was counterbalanced across participants. The location of the payoffs and probabilities (top/bottom) matched the location used during the learning phase.

Risk–reward estimation task (test phase). To test the extent to which (individual) participants had learned about condition-dependent risk–reward relationships, we used payoffs as cues and later asked participants for their estimates of the associated probabilities. As cues, we drew 10 random payoffs (range E$1–1000). The payoffs were identical across conditions.

Recognition (test phase). Finally, to test whether participants recognized specific gambles that did not fit the risk–reward structure of a condition (incoherent gambles "off" the slope, see Figure 2), we asked participants whether they recognized (yes or no) gambles from the learning phase. The recognition task included (1) certainty-effect gambles as a particular case of exemplars that people may recall particularly well, (2) eight environment gambles from the learning phase as a subsample of exemplars that people may have encoded during learning, (3) eight environment gambles that did not appear in the learning phase (but matched the gamble structure of the condition), and (4) eight environment gambles that appeared in the other condition (thus did not match the gamble structure of the condition). This resulted in 32 cued-recognition trials (16 targets, 16 foils; see triangles in Figure 2).

Procedures. Having given signed informed consent, participants were randomly assigned to either the negative or the uncorrelated condition. Participants were only told that they would be asked to make a series of choices between monetary gambles in the first part of the experiment, and that there would then be some additional questions. All experiments were coded in PsychoPy (Peirce, 2007).

In each trial of the learning phase, participants saw a fixation cross (for 500 ms) before making a choice between two gambles. The chosen option was highlighted for 500 ms (by a red rectangle around the gamble). Participants took self-paced breaks after blocks of 30 trials. Gambles were presented in random order. The gambles common to both conditions were randomly interspersed after 50 condition-dependent learning trials.

To link the three tasks of the test phase with the learning phase, we told participants that they would see gambles that were structured similarly to the gambles they had experienced previously, and asked them to think back to these gambles when completing the task given. The order of tasks in the test phase was counterbalanced, with one constraint: Participants always completed the probability estimation task *after* the decision under uncertainty trials to minimize experimental demand effects in the choice task (i.e., prompting participants to infer probabilities from payoff magnitudes).

At the end of the experiment, we played out the chosen option of 20 randomly drawn trials of the learning phase. Bonuses (between 1.92 € and 7.74 €, with E$1000 = 1 €) were added to the regular payment of 10 €/hour.

Analyses. In all experiments, we used a Bayesian approach to data analysis (Kruschke, 2014). Specifically, we applied Bayesian Generalized Linear Mixed Models using Stan in R for regression analyses with the rstanarm package (Stan Development Team, 2016). Unless otherwise noted, we entered participant as a grouping factor to account for individual variation beyond condition-dependent effects. Choice data were analyzed using logistic regressions; estimation data (restricted between [0,1]) were modeled after response data had been transformed to a logit scale. When plotting the posterior-predictive fits of the statistical model, we back-transformed the estimates using the inverse logit. We ran three chains using a Markov Chain Monte Carlo sampler to draw from posterior distributions of parameters. Depending on model complexity, we ran 10,000–30,000 samples per chain (to ensure an effective sample size of $>$ 10,000 for each regressor) and set a burn-in of 500 samples. We investigated (convergence of) our posteriors through visual inspection and the Gelman–Rubin statistic (Gelman & Rubin, 1992). In general, we report the mean of the posterior distribution of the parameter or statistic of interest and two-sided 95% equal tail credible intervals (CI) around each value. Our focus is on estimating the effects of particular conditions and our analyses reflect this goal; in comparing the conditions, however, the crucial issue was whether the credible values included 0 or not.

Results

Decisions under risk (learning phase). We examined choices in the learning phase to see how different risk–reward environments impacted decision making under risk. Four participants chose a stochastically dominated option once, all in the negative condition. The differences in expected values in these trials were small (EV_{abs} = E$60 and E$5, $EV_{\%} = 6.0\%$ and 0.5%) and thus potentially hard to detect. We therefore included these participants' data in further analyses.

Choices between gambles were consistent with standard theories of choice: Participants chose the higher expected value gamble in 79% of all trials ($OR = 5.38$, $b = 1.68$, CI = $[1.48, 1.89]$), and this preference did not differ between the environments ($OR = .82$, $b_{\text{negative}} = -0.19$, CI = $[-0.68, 0.30]$). On a trial-by-trial level, choices depended on how dissimilar the EVs of the two gambles were, with larger EV differences leading to more EV-maximizing choices ($b_{\text{EV}} = .008$, CI = $[.007, .009]$, in a logistic regression with EV differences, higher EV, and condition as predictors, and participant as a grouping factor). This pattern of results persisted when we compared choices in the subset of common gambles only. This finding was contrary to our predictions; we had expected to find systematic differences between environments, particularly for gamble problems aiming to test the certainty effect (see Supplemental Material for details). In sum, we did not find evidence that manipulated risk–reward structures systematically impacted decision-making during the learning phase.

Decisions under uncertainty (test phase). Did payoff levels shape preferences depending on the risk–reward structure experienced? In decisions under uncertainty, participants in both conditions preferred the sure over the uncertain option ($M_{\text{uncertain}} = .21$, $b = 1.89$, CI = $[1.44, 2.40]$). However, as predicted, the strength of preference depended on the learned risk–reward environment and the payoff magnitude offered in the gambles (Figure 3A). Specifically, participants in the negative condition chose the gamble more for low payoffs and less for high payoffs, in contrast to the uncorrelated condition ($b = 1.99$, CI = $[1.03, 2.97]$, payoff $\times$ condition interaction). These choices are consistent with participants in the negative condition inferring probabilities from payoffs, based on the risk–reward structure experienced. In the uncorrelated condition, participants tended to choose the sure thing irrespective of payoff magnitude ($M_{\text{sure}} = .19$, gray line in Figure 3A). Consistent with the prediction that participants in the uncorrelated condition would not use payoffs to infer probabilities, we found that the choices of these participants were independent of payoff magnitudes ($b = 0.15$, CI = $[-.55, .38]$).

Risk–reward estimation task (test phase). Did inferred probabilities reflect previously learned risk–reward structures? Figure 3B shows participants' estimates of the probability of winning a range of payoffs. A negative risk–reward relationship was observed in both conditions ($b_{\text{negative}} = -.78$, CI = $[-.84, -.72]$, $b_{\text{uncorrelated}} = -0.57$, CI = $[-0.63, -0.51]$), but it was stronger in the negative condition ($b = -0.22$, CI = $[-0.30, -.13]$, condition $\times$ payoff interaction; in a regression with condition, payoff, and condition $\times$ interaction as predictors and participant as a grouping factor, using a normal link function).[4] The results in the uncorrelated condition were unexpected in that the choice data suggest that participants had some awareness that they were not in an environment with a negative risk–reward structure. Participants may have been unsure about which probability to indicate in the probability estimation task or have drawn on their mental models of nonlaboratory risk–reward environments to make their estimates.

[4] Here, we used untransformed estimates for better interpretability of the slope. The results are qualitatively identical to those observed when modeling the data with logit transformed probability estimates—the model plotted in Figure 1B.

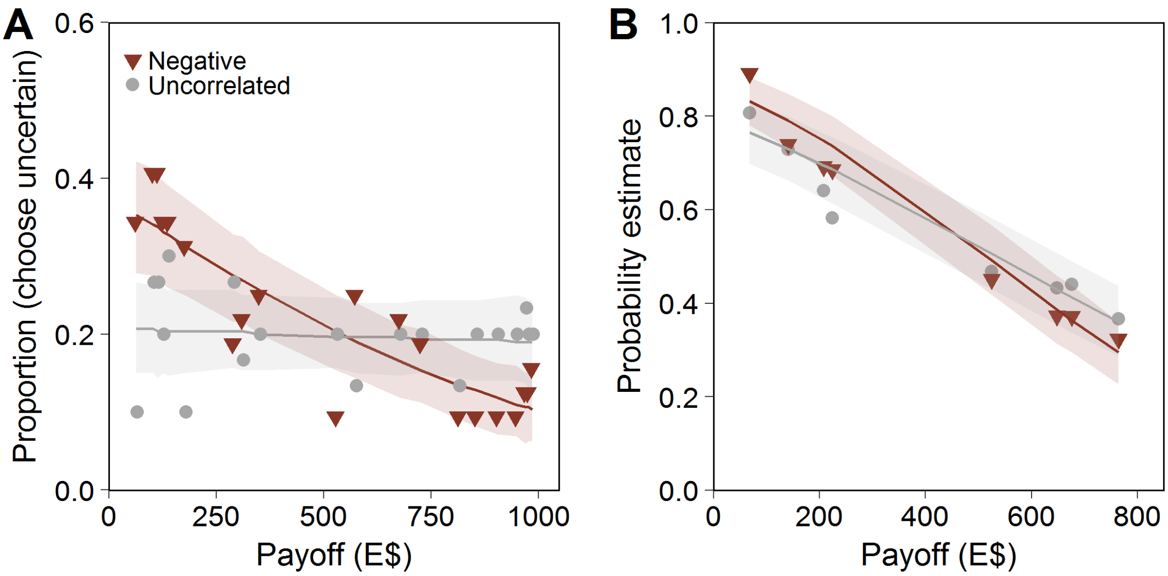

Figure 3. (A) Proportion of times the uncertain option was chosen in the decisions under uncertainty task. Participants in the negative, but not the uncorrelated, condition chose the gamble more for low and less for high payoffs. (B) Average estimated probabilities for each of the possible payoff levels in the probability estimation task. Participants in both conditions estimated an overall negative risk–reward relationship. The solid lines are the posterior predicted means from the respective regression and the ribbons reflect the 95% posterior predictive distribution.

To what degree did individual participants' probability estimates predict their choices in the uncertainty task? To investigate this, we first obtained a (risk–reward) slope for each participant through a random participant term when regressing probability estimates onto payoff magnitudes. This slope served as a measure of participants' judged risk–reward relationship. Steeper slopes indicate a stronger decrease in probability estimates as payoffs increase. We entered these risk–reward slopes in a regression predicting choices from the risk–reward slopes, payoff, environmental condition, and the payoff × condition interaction. The regression showed that steeper risk–reward slopes predicted a stronger tendency to choose the sure thing as payoffs increased, but only for participants in the negative condition (payoff magnitude $\times$ slope, $b = 2.58$, CI $= [0.34, 4.80]$). Individual risk–reward estimates in the uncorrelated condition were not associated with choosing the uncertain option (payoff magnitude $\times$ slope $\times$ uncorrelated interaction, $b = -2.09$, CI $= [-4.79, 0.62]$; modeled in a fixed effects logistic regression, results plotted in Supplement Figure S8A). This result speaks against the possibility that participants in the uncorrelated condition estimated a .5 chance of winning across payoffs and used this subjective estimate across payoffs in decisions under uncertainty. Instead, they estimated an overall negative risk–reward relationship but were averse to uncertainty in their choices across payoffs. This behavior may reflect uncertainty in the estimated chances of winning.

Recognition (test phase). Results from the decisions under uncertainty and probability estimation task both imply that participants were sensitive to the negative risk–reward relationship. How did they learn that relationship? Did they memorize exemplars from the learning phase or abstract the structure as a rule? Results from the gamble recognition task suggest that participants were overall unable to discriminate targets from foils.[5] However, participants did show a bias toward stating that they recognized specific gambles (i.e., saying "Yes"): Of the eight gambles used to study the certainty effect, four fit the negative risk–reward structure (i.e., were "off" the slope) and four did not (see Figure 2A). For instance, whereas winning E$995 would be associated with a low probability in a negative risk–reward environment, the probability in the certainty effect gamble was .8. All of these gambles were used as targets in the recognition task. For gambles that were inconsistent with a negative risk–reward structure (i.e., structured as the bottom left and top right gambles in Figure 2), participants tended to indicate not having seem them previously ($M_{\text{yes}} = .28$, $b = -0.73$, CI $= [-1.25, -0.22]$).

[5] Modeling the data in a signal detection theory framework makes this point clear: The discriminability parameter d' was centered at 0 ($M_{\text{negative}} = 0.0$, CI $= [-0.47, 0.47]$, $M_{\text{uncorrelated}} = 0.00$, $b = 0.00$, CI $= [-0.67, 0.67]$). In addition, participants did not show any systematic response biases in either condition (criterion c, $M_{\text{negative}} = 0.00$, CI $= [-0.27, 0.27]$), $M_{\text{uncorrelated}} = 0.00$, $b = 0.00$, CI $= [-0.41, 0.39]$).

This effect was more pronounced for the negative condition ($M_{\text{yes}} = .17$, $b = -1.14$, CI = $[-1.93, -0.37]$; logistic regression using risk–reward structure, condition, and their interaction as predictors, and participant as a grouping factor).

In sum, the results suggest that the learned risk–reward relationships could not be attributed to participants encoding specific exemplars from the learning phase. Instead, participants in the negative condition may have abstracted a rule that they then used to assess the degree to which the stimuli were consistent with a negative risk–reward relationship. One limitation of the results from the gamble recognition task is that the stimuli set did not include any foils mimicking the structure of the certainty-effect gambles (namely, gambles located at the margins of the payoff–probability space). Instead, all of the extreme gambles were targets. We addressed this issue in the next experiment.

Summary

Experiment 1 exposed participants to either a negative or an uncorrelated risk–reward structure. Our results showed that participants learned whether a risk–reward relationship was present (and negative) or absent (uncorrelated) without any external reinforcement or instruction but via incidental, unsupervised learning.

Moreover, the risk–reward structure impacted preferences in decisions under uncertainty, giving rise to environment-dependent preferences under uncertainty. In the negative risk–reward condition, participants were more likely to prefer the uncertain option with lower payoffs, and their learned risk–reward relationship explained this preference. In the uncorrelated condition, whether participants chose the uncertain alternative was unrelated to payoff magnitudes and estimated risk–reward relationships. Finally, participants (incorrectly) reported not having seen gambles when those gambles were at odds with the negative risk–reward structure, suggesting that they had encoded the overall risk–reward structure as a rule, rather than encoding specific payoff–probability exemplars.

Surprisingly, a majority of participants in the uncorrelated condition estimated an overall negative risk–reward relationship in the estimation task. We offer two possible explanations for this result. One is that that although there was no risk–reward relationship across all gambles in the uncorrelated condition, there was what might be called a local risk–reward relationship within each trial of the learning phase. This is because participants chose between stochastically nondominated options in the learning phase. As a result, for each problem, gamble A will always have a higher payoff but lower probability than gamble B, or vice versa. Thus, participants may have learned a risk–reward relationship from the local as opposed to the global risk–reward relationship. In Experiment 2, we thus modified the learning phase so that a local risk–reward relationship was not present. A second possibility is that participants in the uncorrelated condition responded to the task by harnessing what they know about risk–reward relationships in real-world monetary environments, in which higher payoffs are often less likely to occur (Pleskac & Hertwig, 2014).

Experiment 2: Do People Learn and Exploit a Positive Risk–Reward Relationship?

Experiment 2 sought to replicate and extend the finding that participants are sensitive to risk–reward relationships and harness them in making decisions under uncertainty. To do so, we added an environment with a positive risk–reward structure. Given that this idealistic structure where large payoffs are quite likely to occur is less prevalent outside the lab, it can provide a stronger test of how well participants adapt to different risk–reward structures. In contrast to Experiment 1, we incentivized the decisions under uncertainty task, with the goal to motivate participants to indicate their true preferences and thus better test the environment-dependent preferences observed in Experiment 1. Finally, to create a learning phase without a local risk–reward structure, we asked participants—instead of choosing between two gambles as in Experiment 1—to state the price they would be willing to sell (WTS) a single gamble presented at each trial for.

Based on our findings from Experiment 1, we predicted that participants would learn about risk–reward relationships incidentally while pricing the gambles. The additional positive risk–reward environment affords a stronger, more distinct set of predictions for environment-dependent preferences in decisions under uncertainty. Specifically, we predicted participants in the negative condition should prefer the uncertain option for low payoffs and the sure thing for high payoffs. Conversely, participants in the positive condition should prefer the sure thing for low payoffs and the uncertain option for high payoffs. Additionally, we predicted that choices in the uncorrelated condition would be independent of payoff magnitudes.

Finally, we sought to better understand the degree to which participants encoded the risk–reward relationship as a rule, by modifying the gamble recognition task so that target and foil gambles were structured equally, especially at the four extremes (Figure 4). We hypothesized that participants would not distinguish between targets and foils—which would be difficult to do—but would respond based on the gambles' fit with previously experienced risk–reward structures.

Method

Participants. As in Experiment 1, we aimed for 30 participants per condition. Thus, we recruited 90 participants (53 females, mean age 24.7, $SD = 4.1$, proportion students = .72) from the participant pool at the Max Planck Institute for Human Development. Each participant completed the experiment in exchange for a show-up fee of €10 and a performance-contingent bonus. Participants in Experiment 1 were excluded from the recruitment process.

Decisions under risk (learning phase). The methods were largely the same as in Experiment 1; here, we summarize key differences. We used a larger payoff range (E$1.01–2500, disclosed conversion rate E$2500 = €1). To create the positive risk–reward condition, we took the gambles in the negative condition and reversed the order of

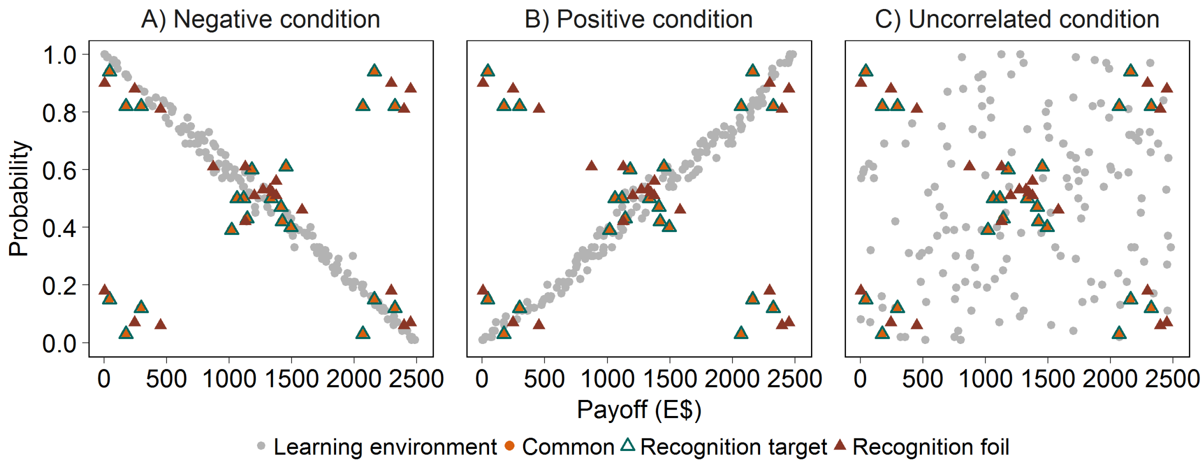

Figure 4. Stimuli used in Experiment 2. The learning phase consisted of 150 condition-dependent gambles, and 22 common gambles (triangles).

probabilities such that the highest probabilities were now associated with the highest payoffs (and vice versa).

As a common set of gambles, we included 10 gambles in the center of the payoff–probability distribution space (intermediate payoffs and probabilities) and 12 gambles at the margins of that space (3 high payoff/high probability, 3 high payoff/low probability, 3 low payoff/high probability, 3 low payoff/low probability; see triangles in Figure 4). Payoffs were random draws between E$0–500 (low) and E$2000–2500 (high). Probabilities were random draws between 0–.2 (low) and .8–1.0 (high). As in Experiment 1, these gambles were used to study whether there were any condition-dependent differences in how participants priced the gambles, while controlling for crucial factors such as EV differences between gambles. In total, these procedures resulted in 172 risky gambles per risk–reward condition while controlling for the marginal distribution of payoffs and probabilities across all three conditions.

Decisions under uncertainty (test phase). For the uncertainty task, we created gamble pairs with low, intermediate, and high payoffs (10 pairs each). As in Experiment 1, the uncertain option's payoff (probability "?") was half as big as the certain option's payoff. In a typical pair, participants chose between a 100% chance of winning E$50 and a "?" chance of winning E$100. We included 30 distractor trials in which the certain option was created by scaling down the uncertain option by a random factor between .1 and .9.

Risk–reward estimation task (test phase). We increased the number of trials such that participants estimated the probabilities associated with 20 payoff magnitudes (range E$1–2500). To investigate how well participants learned the bi-directional relationship between payoffs and probabilities, we additionally asked participants to estimate the payoff associated with a given probability at the end of the experiment. We drew 20 probabilities between 0 and 1 for this task. The results are reported in the Supplemental Material (in short, they precisely mirrored the results of the probability estimation task reported in the paper).

Recognition (test phase). We used the gambles common to both conditions in the learning phase as targets and an equally generated set of gambles as foils (Figure 4, red triangles). Thus, foils were 10 gambles at the center and 12 gambles at the margins of the payoff–probability distribution space (novel random draws based on the recognition gambles procedure). This broader set (relative to Experiment 1) of 22 targets and 22 equally structured foils was used to test whether the risk–reward relationship was learned via exemplars or a rule: If participants had learned the relationship as a rule, they should indicate not having seen gambles that did not fit condition-dependent risk–reward structure (and indicate having seen gambles that did), irrespective of whether those gambles were targets or foils.

Procedure. During the learning phase, participants indicated their WTS for one gamble at a time. They took self-paced breaks after each of five blocks. Common gambles were randomly interspersed after 100 condition-dependent trials. The task was presented as a game show called "Keep or Sell?" ("Behalten oder Verkaufen?"). To motivate participants to indicate their true valuations of a gamble, we implemented a Becker-DeGroot-Marschak auction (Becker et al., 1964). The rules were as follows.

Participants owned the right to play each gamble, which they could sell to the experimenter at a price they determined themselves. Prices were entered with a mouse click on a rating scale (E$0–2500) and confirmed with a click on the value. To incentivize the task, we informed participants that 10 gambles would be randomly selected and played out at the end of the experiment. The experimenter then offered a (computer-generated) buying price between 0 and the maximum payoff from the gamble. If the experimenter's price exceeded the participant's selling price, the participant sold the gamble and earned the buying price. If the participant's selling price exceeded the experimenter's buying price, the gamble was played out (e.g., 50% chance of E$380). The dominant strategy in this task is to price a gamble based on its subjective value: Higher prices can prevent participants from selling unattractive gambles; lower prices can lead to them selling attractive gambles under value. In other words, the prices should approximate participants' subjective certainty equivalents for the gambles.

Participants completed five practice trials to ensure their proper understanding of the WTS measure. If they indicated a selling price that exceeded the maximum payoff from that gamble, participants would see a screen reminding them that (i) they would only receive counteroffers between 0 and the maximum amount to be gained in the gamble, (ii) setting an accurate price would increase the likelihood of good counteroffers, and (iii) good counteroffers would maximize the bonus to be gained from the task. After this feedback, participants set a new price for the same gamble. If they had no more questions, they proceeded to the main part of the task, in which there was no feedback.

The test phase in Experiment 2 was equivalent to that in Experiment 1, except that the decisions under uncertainty task was now incentivized (through five randomly selected trials, using condition-dependent probabilities from the learning phase for the uncertain option). The order of the test phase tasks was again semi-randomized: Decisions under uncertainty preceded the explicit risk–reward probability estimation task, and the risk–reward estimation task (from payoffs) was appended at the end. Across all tasks, the positions of payoffs and probabilities were randomized between subjects; in the decisions under uncertainty task, the positions of gambles were randomized between trials.

Results

Decisions under risk (learning phase). In pricing the gambles, participants integrated payoff and probability information (indicated by a credible payoff × probability interaction, $b = 0.88$, CI $= [0.85, 0.91]$). The prices approximated the gambles' EVs across conditions, with prices in the positive condition deviating slightly more from expected values compared to in the other two conditions (payoff × probability × positive condition, $b = -.07$, CI $= [-.12, -.03]$). However, these differences did not persist when we modeled certainty equivalents given for the subset of gambles common to all conditions (thereby controlling for condition-dependent stimuli features). In addition, there were no differences in participants' subjective evaluations of payoffs and probabilities as modeled by prospect theory (Tversky & Kahneman, 1992) (see Supplemental Material). In sum, and consistent with Experiment 1, participants seemed to evaluate risky gambles in a similar manner across conditions.

Decisions under uncertainty (test phase). How did the experienced risk–reward structures shape participants' preferences under uncertainty? Figure 5A displays the proportion of choices of the uncertain option as a function of the possible payoff level, separately for the three conditions. In general, participants were more likely to choose the certain but smaller payoff option over the uncertain option that offered a larger payoff ($M_{\text{sure}} = .63$, $b = 1.97$, CI $= [1.43, 2.55]$). However, this preference depended on the risk–reward environment to which participants had previously been exposed and on the payoff magnitude offered in the gambles. Consistent with our prediction of environment-dependent preferences, the higher the payoffs, the more often participants in the positive condition chose the gamble ($b_{\text{positive}} = 3.04$, CI $= [2.41, 3.69]$, condition × payoff interaction). When payoffs were high, participants in the positive condition chose the uncertain option in as many as 59% of trials. As Figure 5A shows, the pattern of results was very different for participants in the negative condition, who chose the gamble slightly more for smaller payoffs and less for larger payoffs.

Nevertheless, unlike the results of Experiment 1, the effects in the negative condition were rather small, and the choices were not credibly different from those in the uncorrelated condition ($b_{\text{negative}} = -.22$, CI = $[-.84, .41]$; all effects modeled in a logistic regression with the uncorrelated condition as baseline and participant as a grouping factor). This finding was unexpected: If participants had relied on the learning phase and exclusively used the knowledge they expressed in their probability estimates in the choice task, they should have been much more risk seeking for low payoffs, which they would have learned to be associated with high probabilities. One post-hoc explanation is that participants just tend to reject uncertain options when payoffs are low. However, as Figure 5A shows they still chose the uncertain options offering a low payoff more in the negative risk–reward environment than in the positive one.[6]

Risk–reward estimation task (test phase). Did participants' estimates reflect risk–reward environments from the learning phase? As Figure 5B shows, the probabilities that participants estimated varied as a function of the possible payoffs. Participants in the uncorrelated condition estimated a weak positive relationship ($abs_{\text{slope}} = 0.11$, $b = 0.11$, CI = $[0.06, 0.16]$). Consistent with our predictions, estimates from participants in the other two conditions reflected the specific risk–reward structure to which these participants had previously been exposed. In the negative condition, the estimates showed a negative relationship between payoffs and probabilities ($abs_{\text{slope}} = -.86$, $b = -0.75$, CI = $[-0.82, -0.68]$). In the positive condition, the reverse applied but the slope was much

[6] The choice patterns for the distractor gambles were identical to these results (positive condition choosing gambles more as payoffs increase; negative condition similar to uncorrelated condition); as expected, choices here also largely depended on the difference between certain and uncertain payoffs ($b_{\text{uncertain/certain}} = -8.84$, CI = $[-9.56, -8.16]$).

shallower ($abs_{\text{slope}} = .46$, $b = 0.35$, CI $= [0.28, 0.42]$, in a normal link regression using condition, payoff, and condition × as predictors, and participant as a grouping factor).[7] The shallower slope was partially driven by two participants in the positive condition who estimated a negative risk–reward relationship, perhaps indicative of a possible negative risk–reward relationship 'default'/prior.

To what extent did an individual's probability estimates predict his or her choice in decisions under uncertainty? To address this question, we again obtained a (risk–reward) slope for each participant through a random participant term when regressing probability estimates onto payoff magnitudes. As Figure 6 shows, the majority of slopes (plotted on the x-axis) reflected the condition to which participants had been exposed: the negative condition's slopes fell in the negative range; the positive condition's slopes, in the positive range. We then used the individual slopes to predict choosing the uncertain over the certain option across different payoff magnitudes. The risk–reward slopes predicted payoff-dependent preferences for the uncertain option in the two correlated conditions ($b_{\text{negative}} = 2.01$, CI $= [0.27, 3.76]$; $b_{\text{positive}} = 2.74$, CI $= [1.03, 4.45]$, slope × payoff × condition interaction in a fixed effects model using the uncorrelated condition as baseline). The link between estimates and choices was weaker in the negative condition (red vs. blue slope in Figure 6). As mentioned before, this might be driven in part by an overall tendency to not choose low-payoff uncertain options and thus there is less of an effect of payoff than in the positive condition. Overall, these results suggest that participants inferred probabilities from outcomes in conditions in which a link existed in the gambles to which they had been exposed. As in Experiment 1, risk–reward slopes in the uncorrelated condition did not predict the choice of an uncertain option ($b_{\text{uncorrelated}} = .44$, CI $= [-.56, 1.43]$).

[7] Here, we use untransformed estimates for clarity/better interpretability of the slope. The results were qualitatively identical when we used logit transformed probability estimates to model the data—the model used in Figure 2B.

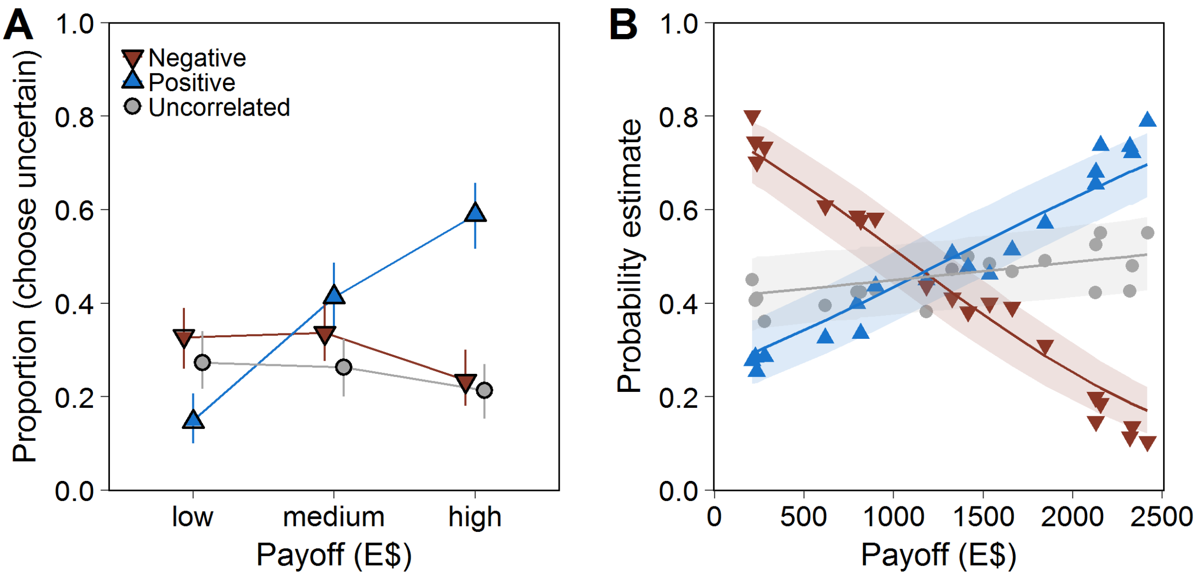

Figure 5. (A) Proportion of times the uncertain option was chosen in the decisions under uncertainty task. Participants in the positive condition chose the gamble more for high and less for low payoffs. Participants in the negative and uncorrelated conditions showed risk-averse behavior, with a low overall proportion of gamble choices. Error bars reflect the 95% posterior predictive distribution; black triangles reflect the mean of the posterior predictive distribution. (B) Average estimated probabilities for each of the possible payoff levels in the probability estimation task. Participants' estimates reflected the risk–reward structures to which they had previously been exposed. The line reflects the mean of the posterior predictive distribution from the linear regression; ribbons reflect the 95% posterior predictive distribution.

Recognition (test phase). Results from the decisions under uncertainty and estimation tasks suggested that participants learned risk–reward structures in the first phase of the experiment. But how did they represent the different structures? Comparison of panels A and B in Figure 7 shows that participants responded similarly when the gambles presented were targets versus foils, implying that they could not discriminate

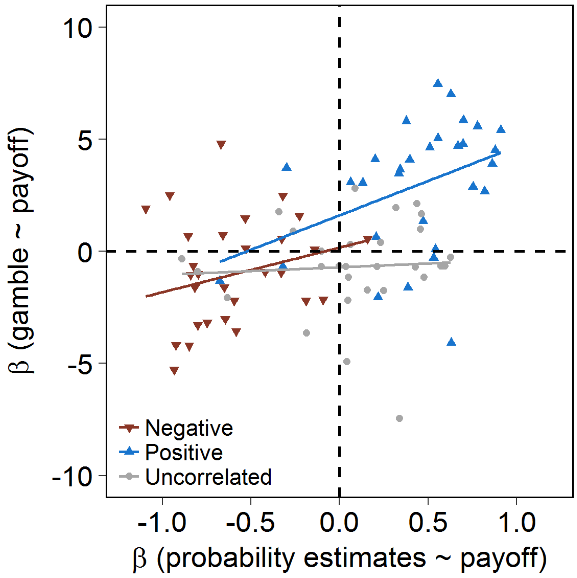

Figure 6. Individual variation in choice of the uncertain alternative (y-axis) based on estimated risk–reward relationships (x-axis). Each data point depicts one participant. The participant-level β was estimated from a Bayesian regression with a random participant intercept for (estimates $\sim$ payoff) and (choice $\sim$ payoff), respectively. Risk–reward estimates in the negative and positive condition, but not in the uncorrelated condition, predicted choice.

between them.[8]

When we broke responses down by whether or not gambles fit a condition's risk–reward structure (Figure 7), the results resembled those of Experiment 1: If gambles did not fit a condition-dependent risk–reward structure, participants indicated that they had not seen them previously, irrespective of whether these gambles were targets or foils. That is, a majority of participants in the negative condition reported not having seen

[8] Signal detection analysis showed that participants did not discriminate between old and new gambles across all three conditions ($b/d'_{\text{uncorrelated}} = 0.21$, CI $= [-0.17, 0.60]$; $d'_{\text{positive}} = -0.16$, $b = -0.37$, CI $= [-0.89, 0.15]$; $d'_{\text{negative}} = -0.05$, $b = -0.26$, CI $= [-0.78, 0.27]$). There were weak, but not credible, biases towards saying 'yes' in the correlated conditions ($b/c_{\text{uncorrelated}} = -0.10$, CI $= [-0.30, 0.08]$), $c_{\text{positive}} = 0.08$, $b = 0.18$, CI $= [-0.09, 0.45]$, $c_{\text{negative}} = 0.03$, $b = 0.13$, CI $= [-0.14, 0.40]$).

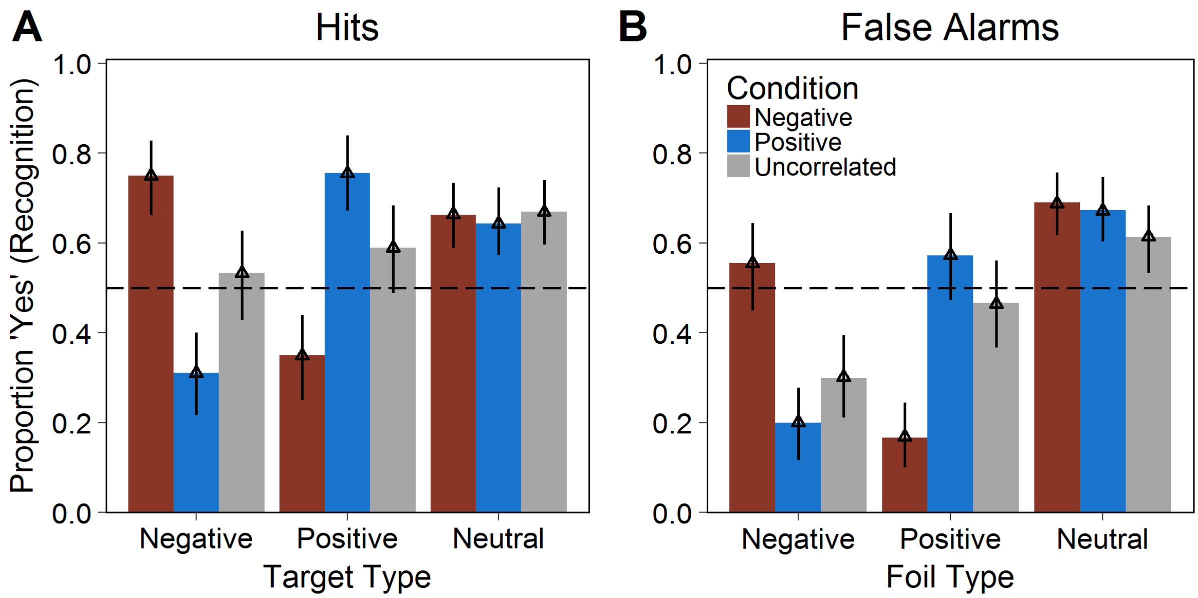

Figure 7. Proportion responding 'yes' to items in the recognition task, by stimuli characteristics and condition. Overall discriminability between targets and foils was low (similar response patterns in panels A and B). Responses depended on whether gambles fit a condition-dependent risk–reward structure. Error bars reflect the 95% posterior predictive distribution; triangles reflect the mean of the posterior predictive distribution.

gambles that were consistent with a positive risk–reward relationship ($M_{\text{yes}} = .26$, $b = -1.40$, CI $= [-1.80, -1.00]$). Conversely, a majority of participants in the positive condition reported not having seen gambles that were consistent with a negative risk–reward relationship ($M_{\text{yes}} = .26$, $b = -.85$, CI $= [-1.25, -.46]$).

What is more, participants were also more likely to report having seen gambles merely because their structure followed a condition's risk–reward structure, again irrespective of whether these gambles were targets or foils (see Figure 7). That is, participants in the negative condition were likely to report having seen a gamble if the gamble was consistent with a negative risk–reward relationship ($M_{\text{yes}} = .65$, $b = .84$, CI $=$ $[.45, 1.23]$), and participants in the positive condition were likely to report having seen a gamble if the gamble was consistent with a positive risk–reward relationship ($M_{\text{yes}} = .66$,

$b = .52$, CI = [.13, .90]; all results from a logistic regression using condition × stimulus type as a predictor, neutral gambles in the uncorrelated condition as baseline and participant as a grouping factor). Responses to neutral gambles were identical across conditions. These gambles were consistent with all risk–reward structures, which might explain why people were equally likely to indicate that they had previously seen them (in all three conditions; $M_{\text{yes}} = .68$, bars on the right in Figure 7A and B). In sum, participants seemed to have learned a risk–reward rule and not to be encoding specific exemplars from the learning phase.

Summary

Experiment 2 substantiated the findings from Experiment 1 that participants learned risk–reward structures in an unsupervised, incidental fashion, and that they subsequently exploited the relationship to make decisions under uncertainty. In particular, we showed that participants learned and used a positive risk–reward relationship, although this structure stands in stark contrast to the negative risk–reward relationship present in many real-world environments. Moreover, in contrast to Experiment 1, where probability estimates in the uncorrelated condition showed a negative association with payoff levels, probability estimates were now independent of payoff levels. One explanation is that this difference may be due to the learning phase, in which participants now evaluated one gamble at a time, removing any 'local' risk–reward structure naturally built into a choice task with nondominated gambles. Probability estimates in the negative condition reflected the structure in the learning phase. This finding suggests that participants used risk–reward relationships to infer probabilities from payoffs—and that the risk–reward relationship learned from pricing gambles dictated the direction of the estimates. Finally, extending the results of Experiment 1, we found further evidence of environment-dependent preferences in decisions under uncertainty. One qualification to this result is that participants in the negative condition were not as keen on choosing the uncertain option for low payoffs as we had expected (despite estimating high probabilities for these payoffs). Finally, the recognition task in Experiment 2 provided further evidence that the risk–reward relationship from the learning phase was represented as a rule rather than in terms of memorized gambles.

Across Experiments 1 and 2, the gambles used in the learning phase presented risks in terms of explicit, single numbers. Outside the laboratory, in contrast, many gambles are about epistemic events, for instance when betting on the outcome of a sporting event (e.g. a soccer match). To gauge their chances of winning for such gambles, people may rely on the prior knowledge they have about these events. Do our results generalize to such choice contexts, in which chances of winning are tied to epistemic events? We turn to this question next.

Experiment 3: Do People Learn About and Exploit Risk–Reward Structures in Bets About Epistemic Events?

In Experiment 3, we examined whether people learn risk–reward relationships when the chances of winning depend on events about which they have some prior knowledge. Specifically, in both the learning and test phase, we used gambles in which winning was tied to the maximum temperature in Berlin on a particular day in 2011 falling within a certain range (e.g., "You win E$500 if the temperature on August 20th, 2011, was between 16 and 25$°C$"). We adapted a procedure from Tversky and Fox (1995) to create different events using different widths and locations of temperature ranges, so that participants should a priori have different subjective probabilities of the events occurring (see also Tversky & Wakker, 1995). As in Experiment 2, participants were asked to state prices for gambles in the learning phase. To create different risk–reward relationships, we determined the probability that a given interval would contain the maximum temperature based on the width of the interval and its proximity to the mean August temperature. We refer to these probabilities as *historical frequencies.* We paired them with payoffs between E$1.01 and

E$2500 to create either a positive or a negative risk–reward relationship. Using these two conditions, we aimed to extend our finding that participants learn risk–reward relationships incidentally from simple monetary gambles to gambles with epistemic events.

Learning about risk–reward relationships from implied subjective beliefs alone may be challenging. Moreover, in some situations, probability estimates about epistemic events are available, such as when an meteorologist shares her belief that an event will occur. Thus, we further differentiated the learning environments, with half the participants being shown only the temperature range of the event but no explicit probability information ('learning under uncertainty') and half additionally being shown the historical frequencies ('learning under risk'). Thus, building on the results of Experiments 1 and 2, by comparing these two sets of conditions, we tested whether explicit probability information is necessary to learn the risk–reward relationship.

In the test phase, we used a similar (though extended) set of tasks as in Experiments 1 and 2 to assess whether participants exploited risk–reward structures in making their decisions. Experiments 1 and 2 suggest that payoff magnitudes can be learned (as expressed in participants' probability estimates) and exploited in choices under uncertainty, in a way that depends on the risk–reward environments to which participants were exposed in the learning phase. In the context of Experiment 3, we assessed the influence of the risk–reward environments somewhat differently. Because the chances of the maximum temperature in Berlin falling within a particular temperature range could be inferred from the range itself, we tested for the effect of the risk–reward environment on probability estimates and choices above and beyond the information provided by the temperature range (historical frequencies). In terms of environment-dependent preferences, this implies for the negative risk–reward environment that the proportion of choices of the uncertain option should increase the lower the payoff, and decrease the higher the payoff—and vice versa for the positive risk–reward environment. A similar pattern should emerge for probability estimates: When relying on a negative risk–reward structure, probability estimates for an event associated with a low payoff should be larger than probability estimates for an (otherwise comparable) event associated with a high payoff. When relying on a positive risk–reward structure, the opposite should happen.

Lastly, we examined whether participants generalized the use of a learned risk–reward relationship to other contexts that they had not learned about and with which they were less familiar. To this end, we added tasks in which participants made decisions under uncertainty and estimated probabilities about the maximum temperature falling in a particular range in Dushanbe, Tajikistan.

Method

Participants. We recruited 200 participants from the participant pool at the Max Planck Institute for Human Development in Berlin (125 females, mean age = 24.45, $SD = 4.3$, proportion students = .84) to take part in the experiment for a 10 € show-up fee and a performance-contingent bonus. Participants in Experiment 1 and 2 were excluded from the recruitment process. Due to the change in design, we expected smaller effect sizes and therefore increased our sample from 30 to 50 participants per condition. Our 200 participants were randomly assigned to one of four conditions. Due to a computer error, the responses from two participants in the uncertainty task were not saved (leaving $N = 198$).

Decisions under uncertainty vs. risk (learning phase). We created bets based on the Berlin weather, an external event that could not be influenced by the experimenter but about which participants should have some beliefs. To this end, we retrieved past weather data on the mean ($M = 22.7°C$) and standard deviation ($SD = 3.2°C$) of the maximum daily temperature in August in Berlin in 2011 from *accuweather.com*. We created 155 temperature ranges of varying width and location on the temperature scale (each August date from 1st–31st was used 5 times). Because the maximum temperatures were approximately normally distributed, we calculated the historical frequency to approximate the probability that the maximum temperature on a given date would fall within the specified interval. We then paired these different events with payoffs between E$1.01 and E$2,500 such that there was either a positive or a negative risk–reward relationship (see Supplemental Material for details). Marginal distributions of payoffs and probabilities were maintained between conditions. We refer to the two conditions *without* additional information about historical frequencies as the "Negative Uncertainty" and "Positive Uncertainty" conditions depending on the underlying risk–reward relationship. In the two other conditions, we added information about historical frequencies to the gamble. We refer to these two conditions as the "Negative Risk" and "Positive Risk" conditions, see footnote for verbal examples.[9]

Decisions under uncertainty (test phase). We created a decisions under uncertainty task in which participants chose between an uncertain option that depended on the weather event occurring ("E$2000 if the maximum temperature was between 23 and 26°C on August 22nd") and a smaller, sure thing ("700 E$ for sure"). We varied the payoffs on two levels, to be either high (E$2000 vs. E$700 for sure) or low (E$100 vs. E$35 for sure). In addition, gambles varied in terms of participants' familiarity with the context. Half of the gambles were about Berlin weather; the other half were about the weather in Dushanbe, Tadjikistan. Being based in Berlin at the time of the experiment, participants were more familiar with Berlin weather.[10] Gambles were presented in blocks; context was counterbalanced between participants.

Subjective probability estimation task (test phase). This task consisted of two parts. First, participants were asked to estimate their subjective probability (0–100%) of winning the gamble (i.e., the event occurring) in the decisions under uncertainty task

[9] Negative Uncertainty: "E$2300 if the maximum temperature was between 13 and 15°C on Aug 29th". Negative Risk: "E$2300 if the maximum temperature was between 13 and 15°C on Aug 29th (p = 3%)." Positive Uncertainty: "E$2300 if the maximum temperature was between 9 and 32°C on Aug 29th." Positive Risk: "E$2300 if the maximum temperature was between 9 and 32°C on Aug 29th (p = 96%)."

[10] This assumption was confirmed in a short post-experiment questionnaire asking participants about their ability to judge Berlin and Dushanbe weather, see Supplemental Material for details.

with payoff information. Our key interest was the degree to which participants used the payoff information in their estimates. Participants were therefore shown the actual gamble (e.g., "E$2000 if the maximum temperature in Berlin was between 23 and 26°C on August 22nd") and asked to judge the probability that they would win. Participants completed this task for both the familiar context (Berlin) and the unfamiliar context (Dushanbe).

In a second part, participants indicated their subjective probability (0–100%) that the maximum temperature on a given day in August would fall in a given temperature range *without payoff information* (e.g., "likelihood the maximum temperature in Berlin was between 23 and 26°C on August 22nd"). The temperature ranges were identical to those used in the decisions under uncertainty task and in the subjective probability estimation task *with* payoff information. Again, we collected estimates for both the familiar context (Berlin) and the unfamiliar context (Dushanbe).

Risk–reward estimation task (test phase). Participants were given payoff information and asked to think back to the gambles they had experienced in the learning phase. They then estimated the likelihood of winning 20 different payoff magnitudes from these gambles in the upcoming bonus trials. This task was used to test whether participants had picked up on the negative versus positive risk–reward structures in the learning phase.

Procedure. Participants were randomly assigned to one of four learning conditions (Negative Risk, Negative Uncertainty, Positive Risk, Positive Uncertainty). They evaluated gambles about the maximum temperature measured on a given day in August 2011 by indicating a WTS for each gamble. The instructions were adapted from the WTS task in Experiment 2. Here, participants were informed that a gamble's value was determined by the extent to which the temperatures were in line with the true temperatures on a given day, and by its possible payoff. The instructions for the risk conditions included an explanation of the historical frequency information added to the gambles: We told participants in the negative and positive risk conditions that the probability was based on typical August temperatures (i.e., “45% is the likelihood that a typical August day will fall in the temperature range given in the bet”). During the learning phase, participants took self-paced breaks between five blocks of 31 pricing trials each. The learning phase was incentivized such that prices from 10 randomly drawn trials were played out according to the Becker-DeGroot-Marschack auction procedure described in Experiment 2, but now the outcome of the gamble was determined by whether the event’s temperature range actually contained the true maximum temperature.

The order of tasks in the test phase was identical across all participants. We randomized the orders of the gamble context (Berlin or Dushanbe) across blocks, and the positions of sure things versus gambles in the decisions under uncertainty task, as well as the position of the payoff amount in the gamble (above or below the event) on the trial level. The decisions under uncertainty task was incentivized such that five randomly selected choices from the task were played out and added to the participant’s total earnings. Finally, we asked people for their subjective estimated ability to judge temperatures in Berlin and Dushanbe (see Supplemental Material).

Results

Decisions under risk vs. uncertainty (learning phase). The WTS prices suggested that participants traded off the payoff and the historical frequencies of events (effect of EV, defined as *historical frequency* × *payoff*: $b = .46$, CI = $[.40, .53]$). As expected, risky gambles that included information about historical frequencies (negative risk and positive risk conditions) were closer to the EVs of the bets than their uncertain counterparts ($b_{\text{negative\_uncertain}} = -.66$, CI = $[-.76, -.57]$, $b_{\text{positive\_uncertain}} = .58$, CI = $[.48, .67]$; 3-way interaction using a gamble’s EV × risk–reward relationship × type of learning, and participant as a grouping factor).

Decisions under uncertainty (test phase). Did the experienced risk–reward relationships shape preferences under uncertainty? We expected this to be the case after "learning under risk" (as in Experiments 1 & 2), in particular, but also (though less strongly) after learning under uncertainty.

For clarity in reporting and interpreting the results, we ran separate analyses for choices about Berlin (Figure 8A, B) and Dushanbe (C, D). In particular, we analyzed condition-dependent choices after controlling for the events' historical frequencies.[11] Figure 8 shows the results of this analysis. Overall, participants were less likely to choose the gamble over the sure outcome for high (E\$2000) than for low payoffs (E\$100) across conditions ($b_{\text{E\$2000}} = -1.23$, CI = $[-1.62, -.85]$). Was this payoff effect moderated by learned risk–reward structures? Indeed, consistent with the risk–reward relationship they had experienced in the learning phase, this payoff effect was smaller for participants who had been exposed to a positive risk–reward relationship under risk (panel A). This effect was driven by participants in the positive condition choosing the gamble less often when the choice was associated with a E\$100 payoff—a payoff that had previously been associated with a low probability ($M_{\text{gamble}} = -.12$, $b_{\text{positive} \times \text{E\$100}} = -.57$, CI = $[-1.08, -.05]$, all results based on a mixed effects logistic regression controlling for historical frequencies, using learning type [risk vs. uncertain] × risk–reward relationship [negative vs. positive] × payoff level as predictors).

Learning under uncertainty did not affect choice in either the familiar context (panel B) or in the unfamiliar context of Dushanbe (overlapping credible intervals across panels C, D). In sum, there is some evidence for environment-dependent preferences, namely when participants were exposed to the risk–reward relationship under risk. Participants in the positive condition became less risk seeking for low payoffs but not more risk seeking for high payoffs, as one would have expected from Experiments 1 and 2.

[11] Participants' choices were well-adjusted to the events' historical frequencies, with an almost linear increase in the proportion of participants choosing the uncertain option as probabilities of winning based on historical frequencies increased ($b_{\text{probability}} = 6.98$, CI = $[6.48, 7.51]$, see Supplemental Material for a detailed plot including posterior predictions across different historical frequencies).

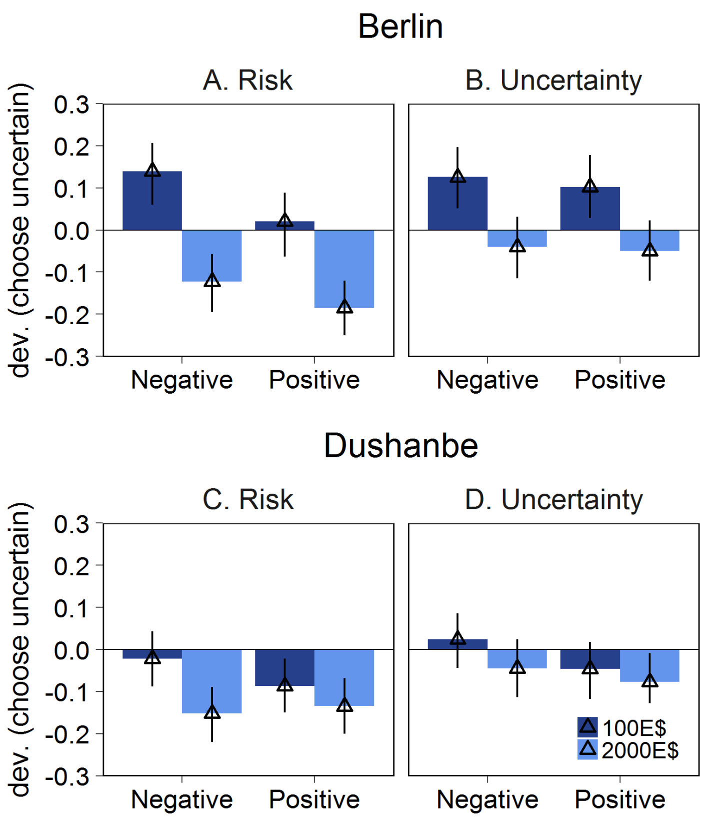

Figure 8. Decisions under uncertainty. Diagrams show how much participants picked the uncertain gamble after controlling for the gamble's historical frequencies (derived from the temperature range). Choice proportions perfectly adjusted to the gambles' historical frequency should have a 0 deviation. Participants in all conditions chose the uncertain option more often when stakes were low (payoff effect E\$100 > E\$2000). Bars and triangles reflect the mean of the posterior predictive choice distribution (controlling for historical frequency); error bars indicate the 95% posterior predictive distribution. Posterior predictions were generated using historical frequencies of .5.

Subjective probability estimation tasks (test phase). In Experiment 3, participants were asked to estimate the chances that a maximum temperature would fall within a given temperature range both within the context of the gamble as a whole (including payoff information associated with the event) and without this payoff information. As we were interested in how the estimates were affected by the risk–reward environments after controlling for the historical frequencies associated with the events, we report deviations from those historical frequencies. For clarity in reporting and interpreting the results, we ran separate regression analyses for estimates about Berlin (Figure 9 A, B) and Dushanbe temperatures (C, D).[12]

Did participants rely on previously experienced risk–reward structures when gauging their chances of winning a bet about the weather? Figure 9 (A, B) shows that participants' subjective estimates were indeed guided by the payoff information. In line with our predictions and Experiments 1 and 2, in the negative conditions (panel A, left bars), subjective probability estimates were lower when temperature ranges were presented in a gamble context that offered a E$2000 payoff ($b = -.10$, CI = $[-.12, -.07]$) than in a gamble context that offered a E$100 payoff.

This payoff effect—a difference in estimates for E$2000 vs. E$100, after learning under risk—was not observed in the positive condition (panel A, right bars, $b = -.04$, CI = $[-.09, .03]$ modeled in a normal link regression using using learning type [risk vs. uncertain] × risk–reward relationship [negative vs. positive] × payoff level as predictors). However, as Figure 9 shows—and contrary to our predictions—the payoff effect did not flip (with higher payoffs leading to a positive deviation and lower payoffs leading to a negative deviation). A bi-product of this was that participants in the risky positive condition ended up with estimates closer to the true historical frequencies.[13] Here, estimates in all three contexts were comparable and relatively close to the historical frequencies of the temperatures learned under risk (panel A). For participants who had learned about risk–reward relationships under uncertainty (panel B), the between-condition effects were comparable

[12] All participants were sensitive to historical frequencies and provided estimates that reflected these frequencies across contexts ($b = .77$, CI = $[.76, .79]$, see Supplemental Material for a detailed plot).

[13] We used unstandardized estimates for clarity. Qualitatively, conclusions remained the same when we used logit transforms of the estimates—the model used in Figure 9.

(larger payoff effect in the negative condition, see panel B), although all estimates now exceeded the historical frequencies of the temperature ranges ($b = .05$, CI = $[.01, .09]$, main effect of uncertainty during learning). Although a similar pattern of results was observed for the unfamiliar context Dushanbe (C, D), there were no credible interaction effects between payoff and learning phase in this context.

Did higher estimated probabilities in this task predict choices in the decisions under uncertainty task? Indeed, we found a link between estimates and choices ($b = 4.00$, CI = $[3.25, 4.78]$, main effect of estimate in a logistic regression using historical frequencies, estimates, and their interaction as predictors, and participant as a grouping factor).

Risk–reward estimation task (test phase). To what extent did probability estimates reflect the experienced risk–reward structures? In the final task, participants were asked how likely they were to win different payoffs in the upcoming bonus trials, based on the learning phase. Figure 10 shows that, as expected, the estimates of participants in the negative condition reflected the risk–reward structure from the learning phase (slope in the negative condition = $b_{\text{negative} \times \text{payoff}} = -.40$, CI = $[-.45, -.35]$). Estimates in the positive conditions were regressive to 50% (slope = .02, $b_{\text{positive} \times \text{payoff}} = .42$, CI = $[.35, .49]$, interaction effect using the negative condition as baseline). Figure 10 (panels A vs. B) also shows that the results were identical for the risky and uncertain learning conditions ($b_{\text{uncertain}} = .02$, CI = $[-.03, .08]$, all estimates modeled in a normal link regression, using learning type (risk vs. uncertain) × risk–reward relationship (negative vs. positive) × payoff as predictors).

Did individual differences in learned risk–reward structures help to predict choices of the uncertain alternative in the decision under uncertainty task? As an index of the learned risk–reward structures, we again estimated a risk–reward slope for each participant from the risk–reward task (payoff-dependent estimates with participant as a grouping factor). In the familiar context (Berlin), there was a weak but not credible association between learned risk–reward relationships and the tendency to choose the higher payoff

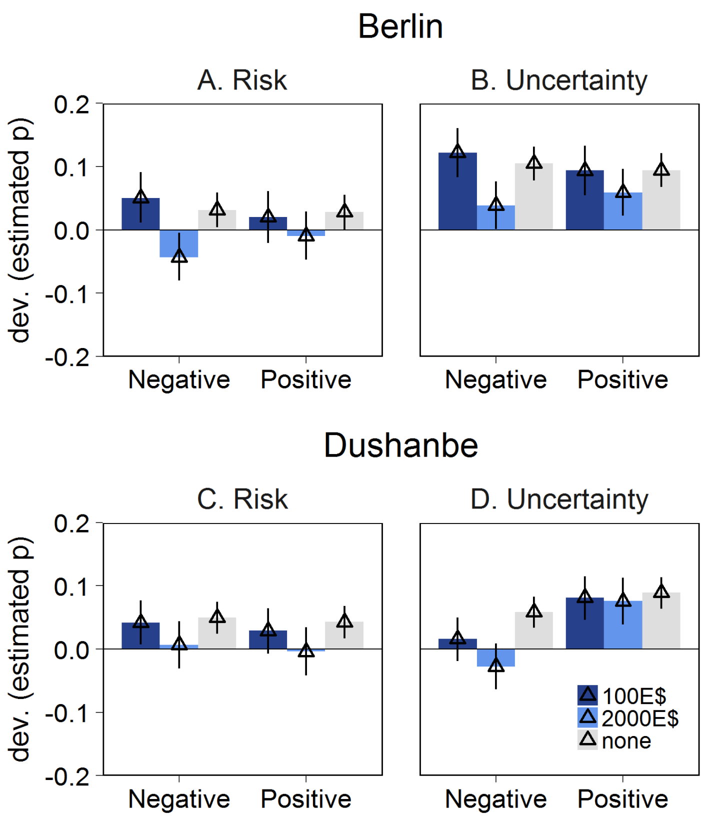

Figure 9. Subjective probability estimation tasks. Plots show deviations of participants' estimates after controlling for the gambles' historical frequencies. Estimates was perfectly in line with the gambles' historical frequency should have a 0 deviation. Participants gave subjective probability estimates of winning a particular temperature bet including a payoff (blue bars) or the probability of a given temperature range alone (gray bars). Bars show posterior mean deviations from historical frequencies; error bars show 95% highest density intervals. Posterior predictions were generated using historical frequencies of .5.

gamble ($b_{\text{positive}} = .64$, CI $= [-.04, 1.33]$, slope $\times$ payoff $\times$ condition interaction in a fixed effects model using the negative condition as baseline, controlling for historical frequencies). Risk–reward estimates did not predict choices in the unfamiliar context Dushanbe ($b_{\text{positive}} = -.13$, CI $= [-.75, .50]$, results plotted in Supplement Figure S8C), suggesting little transfer to less familiar, and unlearned, choice contexts.

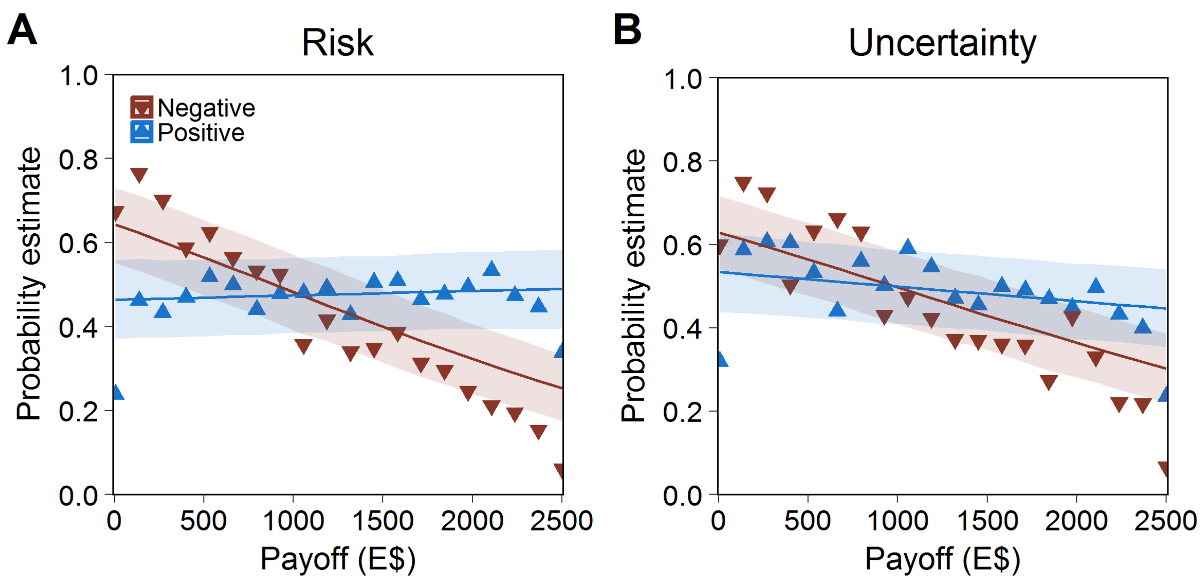

Figure 10. Risk–reward task. Participants were asked to estimate the likelihood of winning different payoffs from learning phase bets. (A) Risk: The learning phase included information about historical frequencies. (B) Uncertain: The learning phase did not include information about historical frequencies. Triangles indicate mean estimates at each payoff level. Lines (ribbons) indicate the mean (95% HDIs) of the posterior predictions.

Summary

Experiment 3 revealed that participants also learned different risk–reward relationships when the probabilities were not explicitly stated as single numbers but expressed in the form of epistemic events. The evidence for learning was more pronounced when the relationship was negative than when it was positive, suggesting that the negative association may have been more in line with participants' initial 'priors'. The learned risk–reward relationship impacted the participants subjective probability estimates about the likelihood of the event occurring. Moreover, preferences in subsequent decisions under uncertainty were to some extent environment-dependent. When participants had explicit probability information available in the learning phase—that is, when they learned under risk, choices were impacted in the low-payoff condition as if participants used both their subjective knowledge about the epistemic events and their knowledge of the risk–reward relationship to estimate subjective probabilities.

We also found that there were limits to the degree to which participants used the risk–reward relationship. Critically, subsequent choices were not impacted when participants learned under uncertainty. In addition, participants did not use the risk–reward relationship in gambles from an unfamiliar domain for which they had not learned the risk–reward relationship.

General Discussion

Ecological structures between risks and rewards that are present in many real-world environments afford decision makers a solution to the problem of unknown probabilities in decisions uncertainty: Decision makers can exploit risk–reward structures to infer probabilities from the magnitude of the payoff itself (Pleskac & Hertwig, 2014). However, such an adaptive, ecological solution to taming uncertainty has specific requirements (Anderson & Schooler, 1991; Brunswik, 1955; Gibson, 1979; Gigerenzer et al., 1991; Marr, 1982; Shepard, 1987; Simon, 1956; Stewart et al., 2006). Here, we investigate two of these requirements: (1) that people are able to extract the environmental structure and (2) that they use the structure adaptively, as the ecological regularities can and do vary across environments (Todd & Gigerenzer, 2007). Our findings from three experiments demonstrate that people do learn risk–reward relationships from the options they experience during preferential choice — without being asked to attend to the structures (incidental unsupervised learning). Moreover, the learned risk–reward relationships guided the direction of estimates and ultimately impacted preferences in decisions made under uncertainty. Next, we discuss our findings with respect to these two requirements and consider their broader implications for adaptive approaches to cognition.

Learning Risk–Reward Structures

Adaptive approaches to cognition seek to understand cognition within the environmental context (Anderson, 1991; Gibson, 1979; Gigerenzer et al., 1999; Marr, 1982; Simon, 1956; Stewart et al., 2006). In the words of Herbert A. Simon 1956, "...we might hope to discover, by a careful examination of some of the fundamental structural characteristics of the environment, some further clues as to the nature of the approximating mechanisms used in decision making" (p. 130). Taking this perspective means it is equally important to identify the ecological structures to which a mind may adapt as it is to establish how the mind comes to terms with those ecological structures (Brunswik & Kamiya, 1953; Simon, 1956).

The risk–reward relationship is an ecological structure that people can use to estimate missing probabilities in decisions under uncertainty (Pleskac & Hertwig, 2014). However, they can only do so if the risk–reward structure has entered the mind. There is good evidence that people are automatic processors of frequency information (a proxy for probabilities) (Hasher & Zacks, 1979; Zacks, 2002), and distributions of payoffs (Brown et al., 2008; Stewart et al., 2006; Olivola & Sagara, 2009; Ungemach et al., 2011). The risk–reward relationship is different in that it is the joint distribution of these dimensions across different gambles, and it is different than a relationship between a cue and decision criterion. Moreover, arguably, learning risk–reward relationships is not a central goal in most decision environments, nor are people explicitly informed of those relationships or learn about them from explicit feedback. Instead, it would seem that, if at all, the risk–reward relationship enters the mind via incidental, unsupervised learning (Brooks, 1978; Dulany et al., 1984; Love, 2002; Nelson, 1984; Ward & Scott, 1987; Wattenmaker, 1991; Whittlesea, 1987). Across three experiments, we showed that participants learn risk–reward relationships from simply expressing their preferences between gambles.

In addition, our results suggest that participants abstracted the relationship as a rule and not via specific exemplars. The strongest support for this conclusion comes from the choice patterns and probability estimates elicited in decisions under uncertainty. Gambles in these tasks did not perfectly map onto learning phase exemplars, yet participants' choices and probability estimates largely resembled the risk–reward rule they had learned previously. Our data cannot, however, pinpoint whether this abstraction occurs during or after encoding (Wattenmaker, 1991). Participants may have used hypotheses about what they know from risk–reward relationships to abstract a rule during encoding (Altmann et al., 1995; Wattenmaker, 1999), or represented the stimuli as exemplars and retrieved a rule from these exemplars as they needed it (Wattenmaker, 1991). A related question is whether risk–reward structures are learned inductively (a negative risk–reward structure and subsequent updating?) or deductively (a tabula rasa each time one enters a new environment).

How much exposure is necessary before people start picking up on a risk–reward relationship in the environment? Our findings point to some general factors that appear to affect how easily risk–reward structures are learned. First, it seems that some risk–reward structures are more difficult to learn than others. Specifically, there was evidence that positive risk–reward structures are more difficult to learn than negative ones: Not all participants picked up on the positive relationship, resulting in weaker positive risk–reward estimates than in the risk–reward structure presented in the learning phase. One possible reason is that people do not come across positive relationships outside the laboratory very often, and thus require more evidence to acquire it. After all, in the real world, there is usually "no free lunch". Second, risk–reward structures seem to be learned more readily with some response types than with others. A comparison of Experiments 1 and 2 suggests that people are more likely to pick up the risk–reward regularity when pricing gambles one-by-one than when choosing between gambles. It is possible that pricing engages deeper processing than choosing (the subjectively better option of two), leading to better encoding of the relationship. Another reason could be that when people choose between two nondominated gambles for all conditions, a 'local' risk–reward relationship is experienced within the choice pair (i.e., a higher payoff is associated with a lower probability relative to the other gamble). A third factor that hampers learning, as Experiment 3 illustrates, is the level of uncertainty in the choices people learn from.

How Risk–Reward Structures Impact Decisions Under Uncertainty

Risks and rewards are the pillars of preference. This makes decision making under uncertainty a vexing problem as one of those pillars—the risks, or probabilities—is missing (Knight, 1921; Luce & Raiffa, 1957). People are commonly thought to deal with this problem by intuiting subjective probabilities from their knowledge and memory (Fox & Tversky, 1998; Tversky & Fox, 1995) or by estimating statistical probabilities from samples of information (Hertwig & Erev, 2009). Our results support another ecologically grounded solution, namely, that people estimate the missing probabilities from their immediate choice environments via their learned risk–reward relationships.

More broadly, these findings fit the general processing assumptions of a risk–reward heuristic. First, the results of the recognition task are consistent with the risk–reward relationship being abstracted as a rule rather than memorized as an exemplar. Second, the subsequent effects of the different risk–reward environments on probability estimates and preferences speak for the subsequent use, or retrieval, of this rule, and against algebraic calculation. Taken together, these properties are consistent with the heuristic use of payoff information to estimate probabilities, rather than with the use of more complex methods. However, our experiments also identified some limitations on the use of the risk–reward relationship as a heuristic. For instance, the results of Experiment 3 demonstrate that other information beyond the payoff information is used to infer probabilities about epistemic events. That is, it is not clear to what degree other information is ignored when inferring probabilities from the payoff information, which is sometimes used as a defining characteristic of heuristics (Gigerenzer & Gaissmaier, 2011; Shah & Oppenheimer, 2008).

Regardless, the exploitation of the environmental structure has some immediate implications. One is that, as we have shown, experienced risk–reward environments can create environment-dependent preferences in decisions under uncertainty. In particular, participants in negative risk–reward environments chose the sure thing more often as payoffs increased, but the opposite occurred for participants in positive risk–reward environments. In uncorrelated environments, preferences were less extreme but still tended to track a negative risk–reward environment, perhaps reflecting the pervasiveness of negative risk–reward environments outside the lab.

Strictly speaking, from a normative perspective, the dependency on payoff information in our experiments would seem to violate the principle of description invariance, according to which the probability of any given event should be judged to be the same, regardless of the associated payoff. This ecological dependency of preferences brings a new perspective to the proposition that preferences are constructed rather than revealed (Ariely & Norton, 2008; Lichtenstein & Slovic, 2006; Payne et al., 1992; Slovic, 1995). The construction of preferences has typically been understood as the result of people selecting a specific procedure from a larger repertoire of possible strategies to formulate a response (Brandstätter et al., 2006; Pachur et al., 2013; Payne et al., 1993; Tversky et al., 1988), the dynamic nature of information accumulation that adjusts preferences over time (Busemeyer & Townsend, 1993), or the ecological (marginal) distribution of monetary payoffs and probabilities (Birnbaum, 1992; Stewart et al., 2006, 2015; Walasek & Stewart, 2015). Here, we have shown how experiencing different risk–reward environments can result in substantial preference shifts in decisions under uncertainty. Edwards (1954) posited that “if utilities and subjective probabilities are not independent, then there is no hope of predicting risky decisions unless their law of combination is known” (p. 400). Our data suggest the ecological relationship between risks and rewards can provide the foundation for such a law of combination.

The finding of environment-dependent preferences also brings a different perspective to other phenomena. For instance, ambiguity aversion—the avoidance of alternatives with unknown probability information when choosing between two otherwise equivalent options (Ellsberg, 1961)—may not be a bias, but partly due to people's choice environments. That is, people may use the risk–reward heuristic to infer the probability of the payoff in the uncertain (or ambiguous) option. If the probability inferred for the unknown option is lower than that offered by the known option, people appear ambiguity averse. Indeed, consistent with this idea, Pleskac and Hertwig (2014) have shown that observed ambiguity aversion increases as the payoff magnitude in comparable options increases. Our results, in fact, suggest that ambiguity aversion can be shaped by mere exposure to different choice environments.

We should emphasize, however, that this change in preferences is not a fallacy, but an ecologically rational bet on the structure of the envrionment. Such a bet is more accurate than ignoring probability information altogether—for example, by using the principle of indifference and assigning equal probabilities to all outcomes (Keynes, 1921). Moreover, our results speak against overtly optimistic estimates that increase as the payoff increases, as implied by the desirability bias (Bar-Hillel & Budescu, 1995; Edwards, 1954; Krizan & Windschitl, 2007; Sharot, 2011) or the affect heuristic (Slovic & Peters, 2006). Instead, probability estimates were adapted to the environment that participants had learned about. If anything, participants adapted too little to positive risk–reward environments, perhaps due to the strength and pervasiveness of negative risk–reward environments.

Conclusion

People often have to make decisions under uncertainty, when probability information is not explicitly stated. In many natural environments, risks and rewards are systematically correlated. This regularity allows people to infer the probability of a payoff from its magnitude, consistent with the use of a risk–reward heuristic. By adjusting their preferences to the respective risk–reward structure, people often manage to make highly adaptive choices under uncertainty.

References

Altmann, G., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 899–912. doi: 10.1037/0278-7393.21.4.899

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*(3), 409–429. doi: 10.1037/0033-295X.98.3.409

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*(6), 396–408. doi: 10.1111/j.1467-9280.1991.tb00174.x

Ariely, D., & Norton, M. I. (2008). How actions create - not just reveal - preferences. *Trends in Cognitive Sciences*, *12*(1), 13–16. doi: 10.1016/j.tics.2007.10.008

Bar-Hillel, M., & Budescu, D. (1995). The elusive wishful thinking effect. *Thinking & Reasoning*, *1*(1), 71–103. doi: 10.1080/13546789508256906

Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, *9*(3), 226–232. doi: 10.1002/bs.3830090304

Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choice-based certainty equivalents. *Psychological Science*, *3*(5), 310–314. doi: 10.1111/j.1467-9280.1992.tb00679.x

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*(2), 409–432. doi: 10.1093/acprof:oso/9780199744282.003.0007

Brooks, D. N. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–215). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, G. D., Gardner, J., Oswald, A. J., & Qian, J. (2008). Does wage rank affect employees' well-being? *Industrial Relations: A Journal of Economy and Society*, *47*(3), 355–389. doi: 10.1111/j.1468-232X.2008.00525.x

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*(3), 193–217. doi: 10.1037/h0047470

Brunswik, E., & Kamiya, J. (1953). Ecological cue-validity of 'proximity' and of other gestalt factors. *The American Journal of Psychology*, *66*(1), 20–32. doi: 10.2307/1417965

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–459. doi: 10.1037//0033-295X.100.3.432

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*(10), 406–416. doi: 10.1016/s1364-6613(98)01232-7

Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications.* San Diego, CA: Academic Press.

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, *113*(4), 541–555. doi: 10.1037/0096-3445.113.4.541

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, *51*(4), 380–417. doi: 10.1037/h0053870

Edwards, W. (1962). Utility, subjective probability, their interaction, and variance preferences. *Conflict Resolution*, *6*(1), 42–51. doi: 10.1177/002200276200600106

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, *75*(4), 643–669. doi: 10.2307/1884324

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*(2), 107–140.

Fox, C. R., & Clemen, R. T. (2005). Subjective probability assessment in decision analysis: Partition dependence and bias toward the ignorance prior. *Management Science*, *51*(9), 1417–1432. doi: 10.1287/mnsc.1050.0409

Fox, C. R., & Rottenstreich, Y. (2003). Partition priming in judgment under uncertainty. *Psychological Science*, *14*(3), 195–200. doi: 10.1111/1467-9280.02431

Fox, C. R., & Tversky, A. (1998). A belief-based account of decision under uncertainty. *Management Science*, *44*(7), 879–895. doi: 10.1287/mnsc.44.7.879

Frensch, P. A., & Rünger, D. (2003). Implicit learning. *Current Directions in Psychological Science*, *12*(1), 13–18. doi: 10.1111/1467-8721.01213

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–511. doi: 10.1214/ss/1177011136

Gibson, J. (1979). *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482. doi: 10.1146/annurev-psych-120709-145346

Gigerenzer, G., Hertwig, R., & Pachur, T. (2011). *Heuristics: The Foundations of Adaptive Behavior.* New York, NY: Oxford University Press.

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A brunswikian theory of confidence. *Psychological Review*, *98*(4), 506–528. doi: 10.1037//0033-295x.98.4.506

Gigerenzer, G., Todd, P., & the ABC Research Group. (1999). *Simple heuristics that make us smart.* New York, NY: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*(1), 75–90. doi: 10.1037//0033-295x.109.1.75

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*(3), 356–388. doi: 10.1037/0096-3445.108.3.356

Hasher, L., Zacks, R. T., Rose, K. C., & Sanft, H. (1987). Truly incidental encoding of frequency information. *The American Journal of Psychology*, *100*(1), 69–91. doi: 10.2307/1422643

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523. doi: 10.1016/j.tics.2009.09.004

Irwin, F. W. (1953). Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality*, *21*(3), 329–335. doi: 10.1111/j.1467-6494.1953.tb01775.x

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291. doi: 10.2307/1914185

Keynes, J. M. (1921). *A treatise on probability.* London: MacMillan and Co., Ltd.

Knight, F. H. (1921). *Risk, uncertainty, and profit.* New York, NY: Sentry Press.

Krizan, Z., & Windschitl, P. D. (2007). The influence of outcome desirability on optimism. *Psychological Bulletin*, *133*(1), 95–121. doi: 10.1037/0033-2909.133.1.95

Kruschke, J. K. (2014). *Doing bayesian data analysis: A tutorial with R, JAGS, and Stan, 2nd ed.* New York, NY: Academic Press/Elsevier.

Lichtenstein, S., & Slovic, P. (2006). *The construction of preferences.* New York, NY: Cambridge University Press.

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review*, *9*(4), 829–835. doi: 10.3758/BF03196342

Luce, R. D., & Raiffa, H. (1957). *Games and decisions.* New York, NY: Dover Publications, Inc.

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review*, *118*(3), 393–437. doi: 10.1037/a0024143

Marr, D. (1982). *Vision: A computational approach.* San Francisco, CA: W.H. Freeman and Company.

Nelson, D. G. K. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior*, *23*(6), 734–759. doi: 10.1016/S0022-5371(84)90442-0

Newsome, J. (2016, March 7). Man wins $291 million in lottery, brother wins $7. *CNN*. Retrieved from http://edition.cnn.com/2016/03/05/us/brothers-win-lottery/index.html

Olivola, C. Y., & Sagara, N. (2009). Distributions of observed death tolls govern sensitivity to human fatalities. *Proceedings of the National Academy of Sciences*, *106*(52), 22151–22156. doi: 10.1073/pnas.0908980106

Pachur, T., Hertwig, G. G., Ralph, & Brandstätter, E. (2013). Testing process predictions of models of risky choice: a quantitative model comparison approach. *Frontiers in Psychology*, *4*, 646. doi: 10.3389/fpsyg.2013.00646

Pachur, T., Hertwig, R., & Steinmann, F. (2012). How do people judge risks: Availability heuristic, affect heuristic, or both? *Journal of Experimental Psychology: Applied*, *18*(3), 314–330. doi: 10.1037/a0028279

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology*, *43*(1), 87–131. doi: 10.1146/annurev.ps.43.020192.000511

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* New York, NY: Cambridge University Press.

Peirce, J. W. (2007). Psychopy - Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Pleskac, T. J., & Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, *143*(5), 2000–2019. doi: 10.1037/xge0000013

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855–863. doi: 10.1016/S0022-5371(67)80149-X

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*(3), 219–235. doi: 10.1037/0096-3445.118.3.219

Savage, L. J. (1954). *The foundations of statistics.* New York, NY: John Wiley & Sons.

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin*, *115*(2), 163–196. doi: 10.1037/0033-2909.115.2.163

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*, *134*(2), 207–222. doi: 10.1037/0033-2909.134.2.207

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*(3), 367–395. doi: 10.1017/S0140525X00035032

Sharot, T. (2011). The optimism bias. *Current Biology*, *21*(23). doi: 10.1016/j.cub.2011.10.030

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 156–163. doi: 10.1016/S0022-5371(67)80067-7

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*(4820), 1317–23. doi: 10.1126/science.3629243

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138. doi: 10.1037/h0042769

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*(5), 364–371. doi: 10.1037/0003-066X.50.5.364

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, *24*(2), 311–322. doi: 10.1111/j.0272-4332.2004.00433.x

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science*, *15*(6), 322–325. doi: 10.1111/j.1467-8721.2006.00461.x

Stan Development Team. (2016). *RStanArm: Bayesian applied regression modeling via Stan (version 2.9.0-4).* Retrieved from https://cran.r-project.org/web/packages/rstanarm/index.html (R package version 2.13.1)

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26. doi: 10.1016/J.Cogpsych.2005.10.003

Stewart, N., Reimers, S., & Harris, A. J. (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science*, *61*(3), 687–705. doi: 10.1287/mnsc.2013.1853

Todd, P. M., & Gigerenzer, G. (2007). Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science*, *16*(3), 167–171. doi: 10.1111/j.1467-8721.2007.00497.x

Tversky, A., & Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review*, *102*(2), 269–283. doi: 10.1037//0033-295X.102.2.269

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. doi: 10.1007/BF00122574

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*(3), 371–384. doi: 10.1037//0033-295X.95.3.371

Tversky, A., & Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica*, *63*(6), 1255–1280. doi: 10.2307/2171769

Ungemach, C., Stewart, N., & Reimers, S. (2011). How incidental values from the environment affect decisions about money, risk, and delay. *Psychological Science*, *22*(2), 253–260. doi: 10.1177/0956797610396225

Wakker, P. (2010). *Prospect theory for risk and ambiguity.* Cambridge, UK: Cambridge University Press.

Walasek, L., & Stewart, N. (2015). How to make loss aversion disappear and reverse: Tests of the decision by sampling origin of loss aversion. *Journal of Experimental Psychology: General*, *144*(1), 7–11. doi: 10.1037/xge0000039

Ward, T. B., & Scott, J. (1987). Analytic and holistic modes of learning family-resemblance concepts. *Memory & Cognition*, *15*(1), 42–54. doi: 10.3758/BF03197711

Wattenmaker, W. D. (1991). Learning modes, feature correlations, and memory-based categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(5), 908–923. doi: 10.1037//0278-7393.17.5.908

Wattenmaker, W. D. (1999). The influence of prior knowledge in intentional versus incidental concept learning. *Memory & cognition*, *27*(4), 685–698. doi: 10.3758/BF03211562

Whittlesea, B. W. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(1), 3–17. doi: 10.1037/0278-7393.13.1.3

Windschitl, P. D., Smith, A. R., Rose, J. P., & Krizan, Z. (2010). The desirability bias in predictions: Going optimistic without leaving realism. *Organizational Behavior and Human Decision Processes*, *111*(1), 33–47. doi: 10.1016/J.Obhdp.2009.08.003

Zacks, R. T. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier & T. Betsch (Eds.), *Frequency processing and cognition* (pp. 21–36). New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780198508632.003.0002