Cognition 175 (2018) 186–200

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Exploiting risk–reward structures in decision making under uncertainty

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

Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

ARTICLE INFO

Keywords:
Decisions under uncertainty
Adaptive cognition
Risk and reward
Incidental learning
Ecological rationality

ABSTRACT

People often have to make decisions under uncertainty—that is, in situations where the probabilities of obtaining a payoff are unknown or at least difficult to ascertain. One solution to this problem is to infer the probability from the magnitude of the potential payoff and thus exploit the inverse relationship between payoffs and probabilities that occurs in many domains in the environment. 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 experiments (N = 352), we found that participants can learn risk–reward relationships from being exposed to choice environments with a negative, positive, or uncorrelated risk–reward relationship. They were able to learn 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 often exploited the learned association by inferring probabilities from the magnitudes of the payoffs. This inference systematically influenced their preferences under uncertainty: Participants who had been exposed to a negative risk–reward relationship tended to prefer 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.

1. Introduction

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 winning this kind of sum 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 associate the highest payoff with the lowest probability? Here, we argue that the key to understanding how the mind generates such estimates lies not within the mind alone, but how the mind is adapted to its environmental context (Anderson, 1991; Gibson, 1979; Gigerenzer, Hertwig, & Pachur, 2011; Marr, 1982; Perkovic & Orquin, 2017; Shepard, 1987; Simon, 1956; Stewart, Chater, & Brown, 2006).

Beyond the lottery, risks and rewards, or payoffs and probabilities, are linked in many choice environments. 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, the strength of the relationship also varies across different domains. Monetary gambles in casinos, for instance, show 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 is less strong. Moreover, a risk–reward relationship is not always given. For instance, no relationship between risk and reward is to be expected in newly forming markets, that have not yet reached an equilibrium (Pleskac & Hertwig, 2014).

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): Risk–reward structures can be exploited 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). Pleskac and Hertwig (2014) offered participants a gamble that gave them a chance to win $$x$ at the cost of $2, and asked

* Corresponding author at: Max Planck Institute for Human Development, Center for Adaptive Rationality, Lentzeallee 94, 14195 Berlin, Germany.
E-mail address: leuker@mpib-berlin.mpg.de (C. Leuker).

https://doi.org/10.1016/j.cognition.2018.02.019
Received 7 November 2017; Received in revised form 16 February 2018; Accepted 16 February 2018

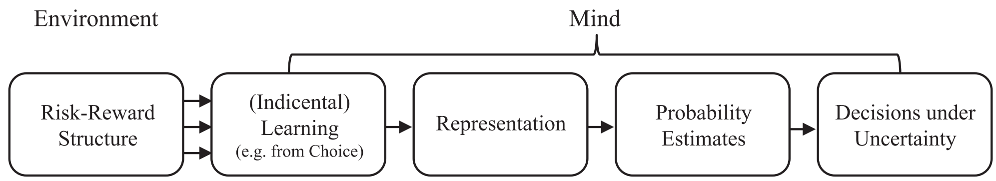

Fig. 1. Summary of the assumed relationships among risk–reward structures in the world and how they ultimately shape preferences under uncertainty.

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. That is, participants inferred the probabilities to be inversely related to the magnitude of the payoff. Moreover, the estimates ultimately influenced what participants chose.

Inferring a probability 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*. Here, we investigate two of its requirements: First, the mind has to be sufficiently sensitive to the relationship between the key variables in an environment (Brunswik, 1955; Gigerenzer, Hoffrage, & Kleinbolting, 1991; Gibson, 1979; Marr, 1982; Simon, 1956; Stewart et al., 2006) or even mirror the relationship from the environment (Anderson & Schooler, 1991; Shepard, 1967, 1987). Second, people should be willing to harness the structure flexibly, as the ecological regularity varies across environments (Todd & Gigerenzer, 2007). That is, there should be a link between the estimates people give and an environments' risk–reward structure. This link also means that, for instance, people should withhold from 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.

Fig. 1 provides an overview of the assumed relationships between risk–reward structures and choice that we take in this paper. Next, we develop our hypotheses in more detail, before reporting three experiments to test them.

1.1. How can people learn risk–reward structures?

In most domains, people are not explicitly told about the presence and/or direction of a risk–reward relationship. They also often do not have the luxury to learn about the relationship 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. In other words, the risk–reward relationship would seem to be learned in an unsupervised manner (without corrective feedback; Love, 2002), and incidentally (when learning is not the primary objective; Brooks, 1978; Dulany, Carlson, & Dewey, 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 learning the frequencies of events, even when that is not their central task (Hasher & Zacks, 1979; Hasher, Zacks, Rose, & Sanft, 1987; Zacks, 2002). People also appear to encode the prices of goods and to use those prices later to evaluate the subjective worth of new values (Brown, Gardner, Oswald, & Qian, 2008; Olivola & Sagara, 2009; Stewart et al., 2006; Ungemach, Stewart, & Reimers, 2011), or use marginal distributions of either payoffs or probabilities in subjective evaluations thereof (Stewart, Reimers, & Harris, 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). It is well known that people can learn associations between two variables (e.g., between a cue and a criterion, see Cooksey, 1996), and sometimes fairly quickly (Kareev, 2000; but see Anderson, Doherty, Berg, & Friedrich, 2005). It is not known whether these findings extend to preferential choice in general; and (maybe even more importantly) to what extent people can learn that there is *no correlation* in their environment, as people may be biased to detecting structures where there are none (Langer, 1975; Olivola & Oppenheimer, 2008).

To test people's ability to learn a risk–reward relationship in an unsupervised, incidental manner, we created a learning phase in which participants encountered gambles where payoffs and probabilities were negatively correlated, positively correlated, or uncorrelated. Across experiments, we tested participants' ability to learn from different types of gambles: 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 stating the price for which they would be willing to sell individual gambles for (Experiments 2 and 3).

Finally, we 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 gamble 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 foils).

1.2. (How) are different risk–reward structures exploited in decisions under uncertainty?

If risk–reward structures are used in decisions under uncertainty to infer the values of missing probabilities, then this can give rise to *environment–dependent preferences*. To see this, consider an environment with a negative risk–reward relationship where high payoffs are unlikely. Someone exposed to this environment is offered a choice between an uncertain gamble with a very high payoff or a smaller, say half-as-large, certain payoff. He or she should prefer the certain payoff (i.e., the sure thing). This is because, according to the risk–reward heuristic, he or she will estimate the chances of obtaining the high uncertain payoff to be quite low and as a result the sure outcome (x) will outweigh the uncertain outcome (y) multiplied by its inferred probability ($x > p_{inferred} \times y$). 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. He or she may adhere to the principle of indifference, assign a

[1] One might also classify this as a case of implicit learning (see, e.g., Cleeremans, Destrebecqz, & Boyer, 1998; Frensch & Rünger, 2003; Reber, 1967; Reber, 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).