Title: Evolution of risk attitudes in the population

Author(s): Erdem Pulcu\textsuperscript{1,*}

Affiliations:
\textsuperscript{1}University of Oxford, Department of Psychiatry, Computational Psychiatry Lab, OX3 7JX UK

*Correspondence to: Dr. Erdem Pulcu, erdem.pulcu@psych.ox.ac.uk
Abstract:
We are living in an uncertain and dynamically changing world. Under uncertainty, perception of risk is an important factor for value-based decision-making and it is directly linked to the survival of species. However, how evolutionary selection pressures might have shaped risk attitudes in the population received almost no attention. Here, we demonstrate that fitness associated with different risk attitudes are influenced by value properties of the environment; as well as the characteristics and the density of competitors in the population. Although the risk neutrality is regarded as the optimal policy, in rapidly changing environments it cannot recover from a marked population density disadvantage. We show that risk-seeking strategies have better fitness when there is: (i) large number of competing strategies in the population, and (ii) when the expected value difference between available options are relatively small. In tandem, the present results may be important for understanding people’s decision-making strategies volatile financial markets in terms of an environmental adaptation.

One Sentence Summary: The present work investigates how natural selection may favour risk-seeking strategies in dynamically changing environments with large number of competitors.

Keywords: Risk perception, value-based decision making, probability weighting, evolution, computational modelling
Main Text:

We are living in an uncertain and an ever-changing world, where our decisions are guided by our expectations of their outcomes. In dynamically changing environments perception of risk associated with the probability of desirable or aversive outcomes is an important factor for modelling decision-making processes (1). Optimal decision-making under uncertainty is a common problem faced by all biological entities in higher classes of the animal taxonomy, and it is crucial for the survival of the species. Consequently, risk perception and decision-making under uncertainty has been studied extensively in economics (2-4), as well as behavioural and neural sciences (5-8); aiming to understand how the brain encodes relevant information from the environment to resolve uncertainty, and makes decisions between the available options. Despite its significance for understanding value-based decision-making, evolution of risk attitudes in the population competing for fitness in dynamically changing environments so far received almost no empirical attention (9, 10).

Here, we addressed this knowledge gap by bridging stochastic choice and stochastic population models in an evolutionary framework; and quantitatively analysed the fitness trajectories of different risk-preferences competing in the population. Our macroscopic approach is important not only because the global financial markets, where millions of traders interact every day, remain just as volatile as the physical environment of the Prehistoric times; but also to understand the evolutionary biological roots of risk-preferences in the population.

Optimal strategy in the deterministic choice model

In order to study how evolution may shape attitudes to risk in the population, we conducted a series of simulated risk decision-making experiments (Fig. 1A), in which there were 4 type of agents defined in terms of their risk taking preferences, namely: risk neutral, risk seeking, risk
averse and risk shifting (see Supplemental Materials and Methods (SMM) for mathematical definitions; and Fig. 1B for their graphical expression); competing in a virtual environment containing 1 million randomly generated options, were rewards are delivered probabilistically (fig. S1). As one might expect, when the agents made decisions in isolation, the risk neutral strategy acquired more resources in an accumulated fashion relative to the other strategies when the expected value difference between the options ($\Delta v$) in the environment varied randomly in a wide (Fig. 1C), or within a limited range ($|\Delta v| < 5$; Fig. 1D). However, these initial simulations in which the agents made decisions in isolation also assumed that agents’ choices are deterministic, such that decision-makers are hypersensitive even to the subtlest changes in the expected value difference between the options, which is not warranted given numerous results from real-life behavioural experiments showing some degree of stochasticity in people's choices(5).
Fig. 1. Decision-making in probabilistic gambles. (A) Schematic diagram of the risk decision-making experiment. (B) Graphical expression of the way the explicit probabilities were modulated under each risk-taking strategy, transforming raw probabilities into subjective probabilities on the y-Axis. (C) The risk neutral strategy acquires the highest accumulated rewards in an environment where the expected value difference between the options are purely randomised or (D) randomised within a restricted range under the deterministic choice regimen. (E) The performances of the strategies with risk aversion are negatively affected under the stochastic choice model ($\beta = 0.8$). (F) While using the risk neutral strategy as a template, assigning higher values to the $\beta$ coefficient in the stochastic choice model shows that accumulated rewards gradually saturate for values of the $\beta \geq 1.6$.

Optimal strategy in the stochastic choice model

In mathematical models of decision-making, the degree of stochasticity is defined by an inverse temperature term ($\beta$) adopted from thermodynamics (also see SMM). Assigning a moderate value to the $\beta$ coefficient, which modulates the subjective value difference between the options (
\( \Delta \tilde{v} \) in a softmax function (which in return generates the choice probabilities of each of the available options (11)): 

\[
q_L = \frac{1}{1 + \exp(-\beta(\Delta \tilde{v}))}
\]  

(1)

suggests that stochasticity will particularly have a negative effect on the performance of strategies with an element of risk aversion (Fig. 1E). Furthermore, by using the risk neutral strategy as a behavioural template, we demonstrate that increasing values of the \( \beta \) coefficient quickly saturates the magnitude of accumulated rewards (Fig. 1F), potentially indicating the upper boundary of its evolution.

**Optimal strategy in the stochastic population model**

Following this rather necessary introduction, we progress with a population level of analysis by duplicating the agents from the first stage to build up a mixed, model society (N=4x10^4) in which each of the 4 different risk taking strategies occupied an equal population density. Here, we linked the individual stochastic-choice model with an evolutionary dynamic computational model (i.e. stochastic population model) at the expected value (\( \Delta v \)) and the choice probability (\( q_L \)) levels (see SMM for the full mathematical description of the models); making it possible to compute the expected random fitness (\( F_A \)) for any of these aforementioned strategies competing against each other to acquire rewards. We created volatile simulation environments by segmenting the original 1 million gambles into 10,000 evolutionary time courses each running for 100 generations, where the expected value difference between the options changed randomly from one generation to the next. The reward magnitudes in each probabilistic gamble corresponded to the amount of resources which can be acquired from the physical environment during the course of one generation on the simulation timeline (Fig. 2A). Once the expected
random fitness \( F_A \) of the competing strategies are computed, it is straightforward to model the local process of the co-evolution by natural selection, as previously proposed by Traulsen et al(12). Natural selection is implemented in terms of bidirectional transition rates between the groups \( r_{A\rightarrow B} \) from one generation to the next and these are based on between-group differences in expected random fitness. In contrast to the results of the individual choice models where agents make decisions in isolation, this methodology reveals that increasing values of the \( \beta \) coefficient enhances the performance of the risk neutral strategy whereby it acquires higher population density at time \( t = 100 \) (the time point where each simulation ended (Fig. 2B~F)), as well as improving the overall fitness of the population (Fig. 2G). Here, it is important to point out that across all the simulations using the stochastic population model reported in this manuscript, the upper boundary of the \( \beta \) coefficient was set to 2.6 which was previously reported by Hsu et al.(5).
**Fig. 2. Evolution of risk attitudes in volatile environments.** (A) Graphical expression of a single simulation environment with respect to expected value difference between the options ($\Delta v$), where each gamble is treated as a generation on the evolutionary timeline. (B-E) Increasing values of the $\beta$ coefficient improves the evolutionary fitness of the risk neutral strategy, while the evolutionary fitness of all the other competitors are negatively affected. (F) Summary of the evolutionary simulations showing the changes in the final normalised population density of each strategy at $t = 100$ against increasing values of the $\beta$ coefficient. (G) Increasing values of the $\beta$ coefficient in the stochastic choice model improves the average magnitude of rewards acquired by the population and it gradually saturates. The error bars denote ±1 SEM across 10,000 simulations.

**Is risk neutrality an evolutionarily stable strategy?**

Although the competition in the population is shown to be the strongest when $\beta = 0.2$, the risk neutral strategy appeared to be the most optimal strategy overall. From an evolutionary fitness
point of view, it is noteworthy that a previous behavioural study reported parameter values for a 2-parameter probability weighting function which would also correspond to risk neutrality in the gains domain (13). However, unlike behavioural studies which focus on participants’ choices in isolation, our population level of analyses show that the performance of the risk neutral strategy also depends on the degree and the characteristics of the volatility in the environment (Fig. 3A~C).
Fig. 3. The nature of environmental volatility influence evolutionary success. The properties of the simulation environment where (A) the risk neutral, (B) the risk seeking and (C) the risk shifting strategies gradually dominate the population, where all the other competitors were driven to extinction. Note that only the first 20 out of 100 generations are visualised for demonstration purposes (i.e. the x-Axes of the line graphs). In lower panels, the evolutionary dynamics of the strategies remain stable after the initial separation of trajectories during the first 20 generations; the strategy with the highest normalised population density in generation 20 gradually dominates the population by generation 100. Here, the population value of the $\beta$ coefficient is set to 0.2.

(D~F) The local process of the co-evolution is decomposed for the risk neutral, risk seeking and risk shifting strategies, respectively; using the environmental information specific for $t = 1$ in each of the environments where they eventually dominate. The numbers in brackets on the decision screen represent reward probabilities, whereas the numbers inside the nodes show choice probabilities under each strategy. The colour of the nodes follow the same legends as the line graphs. (G~I) The local process of the co-evolution is demonstrated with respect to [from top to bottom]: choice probabilities, transitions from and to each group during the first 20 generations where the trajectories of the strategies gradually separate.

Here, by focusing on the first 20 generations where a clear separation occurs between the trajectories of successful risk strategies, we considered three measures to quantify the environmental volatility: magnitude change in the expected value of the environment from one generation to the next; the frequency of the change in the sign of expected value difference (from - to +, or vice versa); and how gradually the expected value difference changed in the environment by checking the correlation coefficient between the vector containing the number of generations and the vector containing expected value differences. Here a highly positive or a highly negative value for the correlation coefficient would mean that environment, although volatile, is changing relatively more gradually from one generation to the next. Across 10,000 simulation environments the risk neutral strategy dominated the population 53.66% of the time, risk seeking strategy 37.58% of the time and the risk shifting strategy 8.76% of the time. It is important to note that across 10,000 different simulation environments, the risk averse strategy is consistently driven to extinction (Fig. 2 B~G). Subsequent analysis suggested that risk shifting strategy prevailed in environments in which the average magnitude of the change in expected value from one generation to the next was highest, and risk neutral strategy when it is lowest (see fig. S2; $F_{2,9999}=20.43$, $p<0.001$, Bonferroni corrected). The environments where different strategies eventually dominated the population, were comparable with respect to other metrics of volatility (all $F_{2,999} < 0.809$; all $p > 0.445$). We further decomposed the local process of the co-
evolution for each of these successful strategies in Fig. 3D–I, demonstrating the divergence of strategies with respect to value properties of the environment. The risk seeking strategy had better fitness relative to other competitors when the expected value difference in the environment varied in a narrow range (\(|\Delta v|<5\)) and \(\beta = 0.2\) (see fig. S3).

In the subsequent step, we investigated how the risk neutral strategy would perform against other transient competitors, those which fall outside of our predetermined/categorical strategies (e.g. risk seeking). In order to address this question in an unbiased way, we preserved the normalised population density of all the strategies at \(t=0\) (\(N_A = 0.25\)). First, we varied the values of the \(\alpha\) and the \(\beta\) coefficients in two probability weighting functions and the sigmoid function simultaneously (see SMM; Eq. 7 and 4, respectively) on a 30x30 numerical grid (i.e. the parameter space) with values ranging from 0 to 3 on each axis. Here, varying the values of the \(\alpha\) coefficient produced strategies with different risk taking preferences against the risk neutral strategy (see SMM for mathematical definitions and fig. S4A for a graphical expression of the probability weighting functions); whereas varying the values of the \(\beta\) coefficient from 0 to 3 covered all possible degrees of stochasticity in agents’ choice probabilities of the competing strategies within this range. The changing values of the \(\beta\) coefficient affected the choice probabilities of all strategies equally; also including the risk neutral strategy. Then, the simulations described above were repeated in the same 10,000 volatile environments to cover all possible combinations of the \(\alpha\) and the \(\beta\) coefficients in this 30x30 parameter space (thus in total performing \(9 \times 10^7\) simulations). The average normalised population density of the risk neutral strategy at each intersection was then converted to a heat map (fig. S4B). This investigation showed that the performance of the risk neutral strategy is sensitive to the changes in the characteristics of the competitors, and there are regions in the parameter space where the risk
neutral strategy loses prevalence from its initial population density, suggesting that absolute risk neutrality cannot be an evolutionarily stable strategy (14), providing additional support for the preceding analysis described above (Fig. 3A–C).

**Could risk neutral strategy emerge from a minority in the population?**

After performing a visual inspection, we selected the intersection of the parameters at $\alpha = 0.8276$ and $\beta = 1.0345$ as a reference point for the next stage of investigations (fig. S4B), aimed at understanding whether the risk neutral strategy could have evolved from a small minority in the population. Although this selection is arbitrary (considering the full size of the parameter space, such an informed selection is necessary for progressive analysis), it serves as a suitable reference point as at this intersection of the $\alpha$ and the $\beta$ coefficients, the risk neutral strategy occupied 51.63% of the population (i.e. average value across 10,000 environments) at the end of the simulation time course despite initially occupying 25% of the population at $t = 0$. This outcome suggests that at this intersection of the parameters, the risk neutrality still remains as the most successful strategy in terms of evolutionary fitness, occupying the majority of the population (>50%). Next, we fixed the $\alpha$ and the $\beta$ coefficients to their values at this intersection and repeated these simulations by assigning a different normalised population density to the risk neutral strategy (values gradually increasing from 0.01 to 0.25) at $t = 0$, with the rest of the three competitors always occupying the remaining of the population with equal population density. Here, we show that in volatile environments the risk neutral strategy cannot recover from a population density disadvantage unless this is almost negligible (Fig. 4A). In order to understand how the fitness of the risk neutral strategy changes in relation to the changes in the characteristics of its competitors, we further zoomed in on to only one of the outcomes of these set of simulations where the risk neutral strategy lost prevalence from its initial population.
density, but remained in the population at the end of the simulation time course in a competitive manner (Fig. 4B). Again, although this selection is arbitrary, it serves as a useful reference point to understand the nature of these competitive interactions; as the risk neutral strategy losing prevalence from its initial population density while not being extinct, suggests that there must be a strong competition between the strategies present in the population. It is important to point out that this informed selection is somewhat necessary due to a high number of output combinations resulting from the interaction between initial population densities and the values of the $\alpha$ and the $\beta$ coefficients (i.e. the grid size of the parameter space multiplied by all possible values which can be assigned to the initial population density). Here, under these constraints, we showed that the landscape representing the outcome of these evolutionary competitions is highly diverse (with emerging plateaus and valleys in the 3D graphical representation; Fig 4B), suggesting that subtle changes in the environmental conditions (including the characteristics of the competitors) may produce quite distinct evolutionary outcomes for the risk neutral strategy, highlighting a degree of randomness in these evolutionary processes occurring within the population (Fig. 4B-C). It is important to clarify that; this is due to the degree of stochasticity in agents’ choice probabilities (i.e. the values of the $\beta$ coefficient) interacting differently with the competitors’ risk-preferences (i.e. the values of the $\alpha$ coefficient) influencing the overall fitness of the risk neutral agents (Fig. 4B).
Fig. 4. The impact of initial population density on evolutionary success of different attitudes to risk. (A) The evolutionary fitness of the risk neutral strategy is influenced by its initial population density at $t=0$. Assigning different population densities to the risk neutral strategy shows that in volatile environments the risk neutral strategy cannot recover from any marked population density disadvantage. The inset figure shows the graphical expression of the risk-taking strategies in the simulation environment: for the value of the $\alpha$ coefficient in both 2-parameter and $\log_2$ probability weighting functions expressed in the figure, in total there are 3 seemingly risk neutral strategies with slight variations in the way objective probabilities were modulated and one with a marked risk averse preference. The dashed red square refers to the final population density of the risk neutral strategy for the combination of the $\alpha$ and the $\beta$ coefficients shown in fig S3B, where it is marked with an asterix sign inside a dashed red square. The dashed blue square shows the reduction in the final population density of the risk neutral strategy due a subtle disadvantage in terms of population density at $t=0$ for the same combinations of the $\alpha$ and the $\beta$ coefficients stated above and in the figure. (B) The landscape of the evolutionary competitions when the initial population density of the risk neutral strategy is set to its value in the dashed blue square in (A). The colour bar shows the average normalised population density of the risk neutral strategy at $t=100$ (mean of 10,000 simulations at each intersection of the $\alpha$ and the $\beta$ coefficients). (C) The heat map showing the normalized population density difference between the risk neutral and the evolutionarily most successful strategy in the population at $t=100$. The colour bar shows the difference in normalised population density. (D) The trajectories of 61 different risk strategies competing in a population. Each trajectory represents the average of 10,000 simulations. The black trajectory is the risk neutral strategy; the red and green trajectories show the least and the most successful strategies, respectively. (E) The graphical expression of the way the objective probabilities were modulated under each of these strategies. Note that here, the most successful strategy (i.e. the green line) is defined by the 2-parameter probability weighting function (Eq. 7a) where the $\alpha$ coefficient is 0.8276, whereas the least successful strategy (i.e. the red line) is defined by the $\log_2$ probability weighting function (Eq. 1b) where the $\alpha$ coefficient is 0.1034 (excluding the strategy which chooses randomly, which is obtained when the $\alpha$ coefficient equals to 0). (F) The histograms showing the distribution of the normalised population density at $t=100$ across 10,000 simulations for each of the strategies in (E): [panels from top to bottom] the most successful; the risk neutral; and the least successful strategy.
Co-evolution of risk-preferences in populations with inherent variability

At the final step, we wanted to investigate the evolution of risk-preferences in a population with a large degree of inherent variability. This scenario is based on a general assumption that behavioural characteristics with large variability (15) observed in the population emerge and carry on existing concurrently, suggesting these may have co-evolutionary dynamics. We constructed such an environment by assigning different values to the $\alpha$ coefficient in both 2-parameter and log$_2$ probability weighting functions, from the parameter space described above. Consequently, in this investigation the population contained 61 different strategies with equal population density at $(t) = 0$: namely, the risk neutral strategy; 30 strategies constructed by assigning different values to the $\alpha$ coefficient in the 2-parameter; and 30 strategies constructed by assigning different values to the $\alpha$ coefficient in the log$_2$ probability weighting functions. We fixed the value of the $\beta$ coefficient in the stochastic choice model to 0.2, as we previously established that this assignment maximised the competition between the strategies (Fig. 2B); so that more information about the nature of evolutionary competitions between coexisting strategies could be obtained. Across 10,000 simulations the risk neutral strategy did not come up as the most optimal strategy in terms of its evolutionary fitness, and it performed relatively poorly compared with a risk seeking strategy (Fig 4D~F). Additionally, we showed that the strategies which subjectively augmented the likelihood of low probability outcomes and attenuated the likelihood of high probability outcomes performed the worst (after excluding the strategy which makes decisions randomly).

Implications for behavioural ecology

The macroscopic/evolutionary approach which is presented here may provide valuable insights for behavioural ecology. Formulating population models of risk-preferences is critical for a
canonical understanding of decision-making processes in predator-prey encounters (16, 17), during foraging considerations (18, 19) and the trade-offs between them (20); all related to evolutionary fitness and natural selection of species. Laboratory studies which could inform development of these population models also suggest that higher order primates are capable of tracking risk associated with pleasant as well as undesirable outcomes (21); with probabilities associated with rewards being encoded in the midbrain dopaminergic (22) and the posterior cingulate neurons (23). Additionally, there is accumulating evidence in favour of variability in risk-preferences across species. For example, bonobos (24) and lemurs (25) show a preference for risk aversion, but rodents (26) and macaques (23, 27) have preference for risky options. A recent study suggests that risk-preferences in animals also adhere to similar nonlinear properties with that of humans (28). Although it is uncertain how well these laboratory findings could represent computations underlying risky decisions in the wild (29), one prediction of our model is that when species are competing to acquire resources in a finite and volatile environment, those with risk-aversion could eventually be selected against.

**Understanding risk aversion in the population and nonlinear modulations of uncertainty**

Inevitably, this prediction would raise a question about the prevalence of risk aversion in the population (4, 30). A considerable number of studies have reported risk-preferences with nonlinear properties, a risk seeking tendency for probabilities approximately lower than 0.34, but a marked risk aversion for probabilities exceeding this threshold (5, 13, 31) (also see SMM and fig. S5). On the other hand, the studies which used a probability weighting function similar to the log2 functional form reported here, focused on how people made value-based decisions while learning the hidden probabilities associated with rewards or punishments by predictive sampling from the environment (1, 32, 33). Arguably, these latter experimental designs have higher
ecological validity in terms of understanding risk-preferences in the population in real life financial decision-making situations, considering that decision-makers do not always have full access to decision variables necessary for computing the expected value difference between the options they face. The probability weighting functions reported by these studies also have nonlinear properties to account for the subjective modulation of risk, but unlike the previous studies mentioned at the beginning, their functional form was mainly expressed in terms of a risk aversion for probabilities lower than 0.5 and a risk prone tendency beyond this cut-off point(1, 32). In the current work, we provide complementary evidence showing that under favourable conditions (Fig. 3C), individuals who utilise a similar probability weighting function to guide their value-based decisions in volatile environments/markets will be the most competitive agents in terms of evolutionary fitness, particularly in environments where the volatility is high in terms of the magnitude of change in resources from one generation to the next (fig. S2). Taken together, log₂ functional form may be another suitable candidate to represent hardwired risk-preferences in humans for everyday financial decisions.

**Potential implications for financial markets**

From a complementary perspective, the findings we present here have important implications for understanding traders’ decisions in global financial markets, which exhibit similar volatile characteristic with that of our simulated environments. It is frequently debated whether risky decisions are among the triggering causes of global financial crises (34) which seem to have shortening cycles. The current investigation shows that risk seeking behaviour could be an evolutionary adaptation among the traders who actualise transactions in financial environments which are both highly competitive and volatile. In this respect, the present results complement the findings of two recent behavioural studies which showed evidence to suggest that observing
others’ risky decisions could influence one’s own risk-preferences in the same direction (35, 36).

Taken together, these individual and population level mechanisms may lead to spread of risk seeking preferences in competitive financial markets.

However, as stated previously, unlike our simulated environments in which the competing agents had full access to the decision variables, the risks associated with returns are mostly hidden in real life and traders have to learn to estimate these from imperfect information(1, 32, 37). The present results suggest that, under these circumstances, the risk seeking strategies would lose their fitness advantage (fig. S6) and such deviations from risk neutrality could increase the stress on global economy until it is discharged in the form of a financial crisis, which happened twice in the last decade.

**Potential implications for understanding behavioural pathologies**

Finally, our macroscopic approach could also inform the evolutionary perspective on psychopathology(38), which posits that clinically debilitating conditions may be associated with fitness advantages. We propose that risk seeking strategies which could be highly adaptive in certain conditions, may contribute to a hardwired, biological vulnerability feature for psychiatric disorders associated with risk and sensation seeking behaviours; such as pathological gambling(39).
References and Notes:

1. T. E. Behrens, M. W. Woolrich, M. E. Walton, M. F. Rushworth, *Nature neuroscience* **10**, 1214 (2007).
2. A. Tversky, C. R. Fox, *Psychological review* **102**, 269 (1995).
3. D. Prelec, *Econometrica*, 497 (1998).
4. D. Kahneman, A. Tversky, *Econometrica: Journal of the Econometric Society*, 263 (1979).
5. M. Hsu, I. Krajbich, C. Zhao, C. F. Camerer, *The Journal of Neuroscience* **29**, 2231 (2009).
6. L. T. Hunt *et al.*, *Nature neuroscience* **15**, 470 (2012).
7. M. Hsu, M. Bhatt, R. Adolphs, D. Tranel, C. F. Camerer, *Science* **310**, 1680 (2005).
8. P. N. Tobler, J. P. O'Doherty, R. J. Dolan, W. Schultz, *Journal of neurophysiology* **97**, 1621 (2007).
9. H.-W. Sinn, *The Geneva Papers on Risk and Insurance Theory* **28**, 87 (2003).
10. L. R. Santos, A. G. Rosati, *Annual review of psychology* **66**, 321 (2015).
11. N. D. Daw, *Decision making, affect, and learning: Attention and performance XXIII* **23**, 1 (2011).
12. A. Traulsen, J. C. Claussen, C. Hauert, *Physical review letters* **95**, 238701 (2005).
13. H. P. Stott, *Journal of Risk and uncertainty* **32**, 101 (2006).
14. J. M. Smith, *Evolution and the Theory of Games*. (Cambridge university press, 1982).
15. L. T. Hunt, *Trends in Cognitive Sciences* **18**, 222 (2014).
16. S. L. Lima, *Trends in Ecology & Evolution* **17**, 70 (2002).
17. M. Hebblewhite, E. Merrill, T. McDonald, *Oikos* **111**, 101 (2005).
18. A. D. Higginson, T. W. Fawcett, P. C. Trimmer, J. M. McNamara, A. I. Houston, *The American Naturalist* **180**, 589 (2012).
19. J. L. Orrock, B. J. Danielson, R. Brinkerhoff, *Behavioral Ecology* **15**, 433 (2004).
20. M. Hebblewhite, E. H. Merrill, *Ecology* **90**, 3445 (2009).
21. V. R. Lakshminarayanan, M. K. Chen, L. R. Santos, *Journal of Experimental Social Psychology* **47**, 689 (2011).
22. C. D. Fiorillo, P. N. Tobler, W. Schultz, *Science* **299**, 1898 (2003).
23. A. N. McCoy, M. L. Platt, *Nature neuroscience* **8**, 1220 (2005).
24. S. R. Heilbronner, A. G. Rosati, J. R. Stevens, B. Hare, M. D. Hauser, *Biology letters* **4**, 246 (2008).
25. E. L. MacLean, T. M. Mandalaywala, E. M. Brannon, *Animal cognition* **15**, 15 (2012).
26. W. Adriani, G. Laviola, *BMC neuroscience* **7**, 52 (2006).
27. B. Y. Hayden, M. L. Platt, *Current Biology* **17**, 49 (2007).
28. W. R. Stauffer, A. Lak, P. Bossaerts, W. Schultz, *The Journal of Neuroscience* **35**, 3146 (2015).
29. F. Paglieri *et al.*, *Frontiers in behavioral neuroscience* **8**, (2014).
30. R. A. Cohn, W. G. Lewellen, R. C. Lease, G. G. Schlarbaum, *The Journal of Finance* **30**, 605 (1975).
31. T. Tanaka, C. F. Camerer, Q. Nguyen, *The American Economic Review* **100**, 557 (2010).
32. S. Suzuki *et al.*, *Neuron* **74**, 1125 (2012).
33. M. Browning, T. E. Behrens, G. Jocham, J. X. O'Reilly, S. J. Bishop, *Nature neuroscience*, (2015).
Acknowledgments:

We would like to thank Dr. Shinsuke Suzuki, Dr. Lalin Anik, Dr. Michael Browning and Prof. Masahiko Haruno for their helpful comments on an earlier version of this paper.
Supplementary Materials:

Supplementary Materials and Methods

Supplementary Methods

Figures S1-S7

References (1-9)

Supplementary Materials Methods:

1. The probabilistic gambles

We have randomly generated $1 \times 10^6$ probabilistic gambles by using MATLAB’s \textit{randsample} function. The reward magnitudes ranged between 10 and 100 with 5 point increments, the probabilities ranged between 0.05 and 0.95 with 0.05 increments. The expected value difference ($\Delta v$) between the left ($\pi_L$) and the right ($\pi_R$) gambles had a mean value 0.02 and standard deviation 30.7 (see fig. S1A; also see below for the notations). Reward magnitudes and probabilities were shown to be decorrelated (see fig. S1B).

2. Definition of different attitudes to risk

The risk averse (RA) and risk seeking (RS) strategies were defined by the 2-parameter probability weighting function\textit{(3)}:

$$\tilde{p} = \exp\left(-\gamma(-\ln p)\delta\right)$$

(1a)

where parameters $\gamma$ and $\delta$ were set to 3 and 1.05 for the RA and 0.5 and 1.05 for the RS strategy, respectively. We also considered a hybrid strategy which has a shifting risk preference (RAS): acting risk averesely for probabilities less than 0.5, and acting in a risk seeking manner for probabilities more than 0.5. It is risk neutral (RN) when the probability is 0.5 and its probability weighting function is defined by the formula:
\[ \hat{p} = 2^{(-\log_2(p)^\alpha)} \]  

where the parameter \( \alpha \) is set to 3. The log\(_2\) functional form always crosses the p/p diagonal at 0.5 and consequently accurately captures the intuition that, psychologically, wide majority of people will have an unbiased perception of the 50/50 odds.

3. Individual stochastic-choice model

Agents adopting any of these aforementioned strategies compute the expected value of a gamble they face accordingly:

\[ \pi = m^* p \cdot p \in \{ p, \tilde{p}_{RS}, \tilde{p}_{RA}, \tilde{p} \} \]  

and make their choices in relation to the subjective value difference between each gamble (i.e. here, the difference between left and right options):

\[ \Delta \tilde{v} = \pi_L - \pi_R \]  

trial-wise stochastic choice probabilities of each option follow Luce’s choice axiom and choice probabilities for each gamble are generated by a sigmoid function:\( (11) \)\:

\[ q_L = 1 / (1 + \exp(-\beta(\Delta v))) \]  

where \( \beta > 0 \), is the inverse temperature term adopted from thermodynamics and it determines the degree of stochasticity in choice probabilities; values of \( \beta \to 0 \) giving way for stochastic choices, and values of \( \beta \to \infty \) leading to deterministic choices.
4. The stochastic population model

After defining the stochastic choice model for the value-based risk decision-making at the single agent level, we constructed the stochastic population model by applying a kinetic Monte Carlo algorithm (40), to define the local process of the evolutionary game (12, 41). At time \( s \in [0, T_{end}] \), we defined the random populations of each of the groups as \( N_A(s), A \in \{RN, RS, RA, RAS\} \).

The local process of the between-group competition (12) is then defined accordingly:

\[
N_A \rightarrow N_A - 1, N_B \rightarrow N_B + 1, \text{ with a rate of } N \cdot r_{A\rightarrow B} \tag{5a}
\]

applicable for all possible combinations of \( A \) and \( B, \in \{RN, RS, RA, RAS\} \).

Here, \( N := N_{RV} + N_{RS} + N_{Ra} + N_{RAS} \) is the constant population size of the system (i.e. a linear Moran process (41)), which is fixed to \( 4 \times 10^4 \) agents with each of the competing groups occupying \( \frac{1}{4} \)th of the population for an unbiased investigation of their evolutionary fitness.

The expected random fitness of any agent in group \( A \) is defined by

\[
f_A := q^A_L \cdot \pi_L + (1 - q^A_L) \cdot \pi_R \tag{5b}
\]

The transition rates between the groups, \( r_{A\rightarrow B} \)'s, are then defined by the formula:

\[
r_{A\rightarrow B} := \frac{1}{2} \frac{N_A}{N} \frac{N_B}{N} (1 + \frac{F_B - F_A}{\Delta F_{max}}) \tag{5c}
\]

In this formulation, \( F_A \) is the random (expected) fitness of the group \( A \):

\[
F_A := f_A \cdot N_A \tag{5d}
\]

It is important to point out once again that the agents’ choice probabilities are based on their subjective value difference (\( \Delta v \)), whereas their expected random fitness is based on the
average expected values of each option ($\pi$) computed under the risk neutral regimen for all types of agents (i.e. how much rewards the proportion of the agents choosing one option should actually expect to receive from the physical environment). $\Delta F_{\text{max}}$ in Eq.(5c) serves as a normalisation constant to make sure that $r_{A \rightarrow B} \geq 0$; such that the transition rates between the groups will always remain positive (41), and it is calculated by the following formula:

$$\Delta F_{\text{max}} := N \cdot (\pi_L + \pi_R)$$  \hspace{1cm} (5d)

considering the full range of the value space of the dynamically changing, volatile environment in any given generation, which is also the limit of the maximum fitness difference which could be observed between any possible competitors in any generation.

Exact group trajectories were generated by the standard kinetic Monte Carlo algorithm proposed by Gillespie (40), whereby the population density of group $A$ at time $(t + 1)$ from time $(t)$ is calculated as follows:

$$N_{A}^{t+1} = (N_{A}^{t} - \sum_{B \in \{N, RS, RA, NL\}} N \cdot r_{A \rightarrow B}^{t} + \sum_{B \in \{N, RS, RA, NL\}} N \cdot r_{B \rightarrow A}^{t})$$ \hspace{1cm} (6)

We simulated between group competitions in different settings (i.e. 4 different $\beta$ values) in 10,000 volatile environments (containing of 100 randomised gambles with different expected value difference ($\Delta v$)). Visual inspection of the Standard Error of Measurement (SEM) margins (i.e. shaded area around the mean trajectories) suggests that the degree of volatility and the behaviour of the strategies across different simulations were mostly comparable. In support for the reliability of the agent-based model we propose here, our recent investigation shows that when competing groups sizes are $\geq 3000$ agents, outcome of the kinetic Monte Carlo simulations converge with the trajectories obtained from solving deterministic, mean-field [differential] replicator equations (42) which define the evolution of the system at the infinite population limit.
5. Generating intermediate risk strategies by varying the values of the $\alpha$ coefficient

In the follow-up analyses to the ones shown in Fig.2 B–E, evolutionary fitness of the risk neutral strategy was tested against different risk strategies which were generated by varying the values of the $\alpha$ coefficient from 0 to 3. Here, the traditional $\gamma$ coefficient was replaced by the $\alpha$ coefficient in the 2-parameter probability weighting function which is used to define the original risk seeking and risk averse strategies, and the $\delta$ coefficient was fixed to its original value (i.e. $\delta = 1.05$). The parameter space of the $\alpha$ coefficient was defined by MATLAB’s linspace function, whereby 30 possible values of the $\alpha$ coefficient between 0 and 3 were generated. Thus, for every value of the $\alpha$ coefficient, the volatile simulation environment contained 3 competitors against the risk neutral strategy, defined by the following equations:

\[ \tilde{p} = \exp^{-(\alpha(-\ln(p)))^\delta} \]  

(7a)

by which the original risk seeking strategy was modified; and

\[ \tilde{p} = \exp^{(-(3-\alpha(-\ln(p)))^\delta)} \]  

(7b)

by which the original risk averse strategy was modified.

The original risk shifting strategy was modified by assigning values to the $\alpha$ coefficient in Eq. (1b) from the same parameter space defined above.

The graphical expressions of these functions are given in fig S4A.

6. Complementary analysis involving Prelec 1-parameter model with human hyper-parameters

We also investigated the fitness of a strategy driven by the 1-parameter probability weighting function, which is in essence a modified version of the 2-parameter probability weighting function(3) which was used to define the risk seeking and the risk averse strategies:

\[ \tilde{p} = \exp^{-(\gamma(-\ln(p)))^\delta} \]  

(8)
where, the 1-parameter probability weighting is obtained by setting the $\gamma$ and the $\delta$ to 1 and 0.74, respectively; as reported by two independent human risk decision-making studies (31, 43). As a comparison, a recent study using the same 1-parameter function in monkeys reported $\delta$ values of 0.31 and 0.47 respectively for two different experimental animals (28).

In this complementary analysis, we first replaced the risk shifting strategy by the 1-parameter function, while keeping all the other strategies the same as the earlier simulations (fig. S5A); and later on included the 1-parameter function with human hyper-parameters as the fifth strategy in the population (fig. S5C).

Please note that in all of the complementary analyses described in this section the value of the $\beta$ coefficient was set to 0.8.

7. Performance of strategies in noisy environments

In another complementary analysis, we investigated the impact of inherent noise in the environment such that even the risk neutral strategy fails to process information related to explicit probabilities optimally. Considering that the probabilities were originally defined within the range $0.05 < p < 0.95$, we ensured that the noise level is always less than the explicit probability information ($\varepsilon_{\text{max}} < p_{\text{min}}$). In this set of simulations, the values of the noise term, $\varepsilon$, were generated by MATLAB’s `rand` function by the explicit operation: $\varepsilon = \text{rand} / 20$.

We considered 4 different scenarios in which the probabilities were: (i) always lower (fig. S6A); (ii) always higher (fig. S6B); (iii) always higher on the left side, but lower on the right side (fig. S6C); and finally by random assignment (iv) low or high on the both the left and the right side (i.e. 2x2 factorial design, fig. S5D) than explicit probabilities due to noise in the environment.

8. Performance of strategies with alternative definitions of risk proneness and risk aversion

As a part of our final complementary analysis, we wanted to re-investigate the competition between our categorical strategies by redefining them in terms of another value function which describes risk seeking and risk averse preferences. Here, we relied on the exponential utility function, which is quite frequently used in the literature, but does not make any claim on
probability modulation under uncertainty. Note that, in this investigation the risk neutral and risk shifting strategies remained in the population and they were preserved as in Eq. 1b and 2, whereas both the risk averse (RA) and risk seeking (RS) strategies were redefined by the following equation:

$$\pi = m^\rho \cdot p, \rho \in \{\rho_{RA}, \rho_{RS}, \rho_{RN}\}$$ (9)

where the parameter $\rho$ defines individuals’ risk preference. We derived the value of $\rho$ from the population average reported by Hsu et al., which suggests risk aversion $\rho_{RA} = 0.576$. Here, the categorical risk seeking strategy was defined by its mirror distance from risk neutrality ($\rho_{RN} = 1$), where $\rho_{RS} = 1 + (1 - \rho_{RA})$, and therefore 1.424. This analysis also suggests that risk seeking strategies, defined by the exponential utility function, may have better evolutionary fitness relative to other competing strategies (fig S7).
Supplementary References:

1. Prelec D. The probability weighting function. Econometrica. 1998; 497-527.
2. Daw ND. Trial-by-trial data analysis using computational models. Decision making, affect, and learning: Attention and performance XXIII. 2011; 23: 1.
3. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. Journal of computational physics. 1976; 22(4): 403-34.
4. Traulsen A, Claussen JC, Hauert C. Coevolutionary dynamics: from finite to infinite populations. Physical review letters. 2005; 95(23): 238701.
5. Bladon AJ, Galla T, McKane AJ. Evolutionary dynamics, intrinsic noise, and cycles of cooperation. Physical Review E. 2010; 81(6): 066122.
6. Traulsen A, Claussen JC, Hauert C. Coevolutionary dynamics in large, but finite populations. Physical Review E. 2006; 74(1): 011901.
7. Wu G, Gonzalez R. Common consequence conditions in decision making under risk. Journal of Risk and uncertainty. 1998; 16(1): 115-39.
8. Tanaka T, Camerer CF, Nguyen Q. Risk and time preferences: linking experimental and household survey data from Vietnam. The American Economic Review. 2010; 100(1): 557-71.
9. Stauffer WR, Lak A, Bossaerts P, Schultz W. Economic choices reveal probability distortion in macaque monkeys. The Journal of Neuroscience. 2015; 35(7): 3146-54.
Supplementary Figures and Legends:

**A**

**Distribution of Value Difference in 1x10^6 Gambles**

- Histogram for the distribution of the expected value difference between the options ($\Delta v$) across 1x10^6 probabilistic gambles.
- The reward magnitudes and probabilities for each option were decorrelated. The colour bar shows the correlation coefficient ($r$).

**B**

- Visual representation of the correlation matrix between options.
- The colour bar indicates the correlation coefficient ($r$) ranging from 0 to 1.

**fig. S1. Properties of the simulation environment.** (A) Histogram for the distribution of the expected value difference between the options ($\Delta v$) across 1x10^6 probabilistic gambles. (B) The reward magnitudes and probabilities for each option were decorrelated. The colour bar shows the correlation coefficient ($r$).
**fig. S2. Metrics of environmental volatility with respect to winning strategies.** In environments where the risk shifting strategy dominated the population, volatility was significantly higher in term of the change in expected value difference from one generation to the next (\( \Delta v \); *** p<0.001, Bonferroni corrected).

With respect to the frequency of sign change and how gradually environments changed down the evolutionary timeline, despite different strategies dominating the population, there were no significant differences between the environments in terms of volatility (n.s: not significant).
fig S3. Evolution of risk preferences in low value difference environment with respect to changing values of the $\beta$ coefficient. (A) In low expected value difference environments where the $\beta$ coefficient in the stochastic choice model is set to 0.2, the risk seeking strategy has better evolutionary fitness. The trajectories show the average of 10,000 simulations. (B) Across the same 10,000 simulation environments, when the value of the $\beta$ coefficient is set to 2.6 (i.e. the population mean reported by Hsu et al. (5)), the risk neutral strategy dominates the population. Note that the risk averse and the risk shifting strategies have overlapping trajectories, suggesting that in low expected value difference environments these strategies eventually make the same decisions.
**fig. S4. Variability in risk preferences with respect to the changing values of the $\alpha$ coefficient. (A)**
The graphical expression of the way the objective probabilities were modulated for each of the strategies which compete against the risk neutral strategy while varying the values of the $\alpha$ coefficient in Equations 1b, 7a and 7b simultaneously. The colour bar shows the values of the $\alpha$ coefficient in the probability weighting functions. **(B)** The heat map shows the average normalised population density of the risk neutral strategy at $t = 100$, averaged across 10,000 simulations. The red dashed square with an asterix shows the reference point for the combination of the $\alpha$ and the $\beta$ coefficients, where the risk neutral strategy occupies more than half of the population at $t = 100$, despite initially occupying 0.25 of the population (as with all the other competitors). The values of the $\alpha$ and the $\beta$ coefficients on this reference point were then used for the analysis described in Fig. 4 and legends.
**fig. S5. The evolutionary fitness of the risk profile based on human hyper-parameters.** (A) Prelec 1-parameter function with specific human hyper-parameters based on previous studies (31, 43) performs better than categorically defined risk seeking and risk averse strategies. (B) The graphical expression of the Prelec 1-parameter function with respect to risk neutrality (i.e. p/p diagonal). (C) The same Prelec 1-parameter function remains competitive even when there are more categorically defined strategies in the population, performing second to best after the risk neutral strategy.
fig. S6. Competitions between the strategies in noisy environments. Despite face value, probabilities had an intrinsic element of noise, where they were actually always (A) lower; (B) higher; (C) higher on the left side but lower the right side than they appear to be. (D) The environment where the noise is randomly assigned from a 2x2 factorial matrix (i.e. combinations of left and right; with low and high). In all cases the risk neutral strategy is the most competitive strategy in the population.
fig. S7. Evolution of risk attitudes when competing risk strategies were defined by the utility function. Redefining the risk seeking and risk averse strategies by the exponential utility function shows that the agents adopting the risk seeking strategy, whose choices are driven by the exponentially increasing utility derived from linearly increasing reward magnitudes, would acquire better evolutionary fitness relative to their competitors, also including those who adopt the risk neutral strategy.