Behavioral/Systems/Cognitive

Neural Correlates of Anticipation Risk Reflect Risk Preferences

Sarah Rudorf,[1] **Kerstin Preuschoff,**[2] **and Bernd Weber**[1,3]

[1]Center for Economics and Neuroscience, University of Bonn, 53127 Bonn, Germany, [2]Social and Neural Systems Laboratory, Department of Economics, University of Zurich, 8006 Zurich, Switzerland, [3]Department of Epileptology, University Hospital Bonn, 53127 Bonn, Germany

Individual risk preferences have a large influence on decisions, such as financial investments, career and health choices, or gambling. Decision making under risk has been studied both behaviorally and on a neural level. It remains unclear, however, how risk attitudes are encoded and integrated with choice. Here, we investigate how risk preferences are reflected in neural regions known to process risk. We collected functional magnetic resonance images of 56 human subjects during a gambling task (Preuschoff et al., 2006). Subjects were grouped into risk averters and risk seekers according to the risk preferences they revealed in a separate lottery task. We found that during the anticipation of high-risk gambles, risk averters show stronger responses in ventral striatum and anterior insula compared to risk seekers. In addition, risk prediction error signals in anterior insula, inferior frontal gyrus, and anterior cingulate indicate that risk averters do not dissociate properly between gambles that are more or less risky than expected. We suggest this may result in a general overestimation of prospective risk and lead to risk avoidance behavior. This is the first study to show that behavioral risk preferences are reflected in the passive evaluation of risky situations. The results have implications on public policies in the financial and health domain.

Introduction

To a great extent, our decision behavior is determined by individual preferences. We all prefer higher rewards over lower rewards, but the picture is less clear when it comes to risk. In decision making under risk, outcomes are not certain but occur with known probabilities. Suppose a decision maker has a choice between a sure payout of €10 and a risky lottery that pays €25 with a probability of 50% and €0 otherwise. Here, the decision maker faces a trade-off between the higher expected reward of the lottery and the lower risk (the variance of outcomes) of the sure payout. This trade-off is captured by economic models of choice under risk. The trade-off may be explicit as in the mean–variance approach (Markowitz, 1952) or implicit as in prospect theory (Tversky and Kahneman, 1992). These models, when applied to choice behavior, reveal a wide range of individual preferences toward risk. These preferences (risk aversion, risk seeking) are reflected in our daily lives and may lead to suboptimal choices in health, investment, and career decisions (Guiso and Paiella, 2005; Dohmen et al., 2011). Although risk preferences are very heterogeneous across domains, they can be well predicted within domains by experimental measures (Weber and Hsee, 1998; Blais and Weber, 2006). In this regard, various disciplines could ben-

efit from the experimental investigation of risk preferences. Understanding how economic risk preferences are integrated in the cognitive mechanism underlying decision making under risk, for instance, may help us to improve investment strategies that not only aim at the highest possible return, but take into account how much risk the investor might tolerate. We therefore seek to understand whether risk preferences are reflected in the neural correlates of risk.

In the example above, risk is present during the choice ("decision risk"), and, if the lottery is chosen, after the choice while the decision maker is awaiting the lottery's outcome ("anticipation risk") (Mohr et al., 2010). For decision risk, risk preferences have been associated with neural responses: risk-seeking people, for instance, are characterized by a lack of prefrontal inhibition of risk-taking behavior (Gianotti et al., 2009). But it remains largely unknown how risk preferences are reflected in the neural correlates of anticipation risk, i.e., during the passive anticipation and the subsequent realization of risky outcomes (Preuschoff et al., 2006). In previous studies we showed that the ventral striatum and anterior insula reflect anticipation risk (Preuschoff et al., 2006; Preuschoff et al., 2008). When the risky gamble is resolved, the deviation of the anticipated risk from the actual outcome variance results in a "risk prediction error." This risk prediction error has been located in the anterior insula and the inferior frontal gyrus (IFG) (Preuschoff et al., 2008). Here we ask whether individual risk preferences are reflected in these representations.

Materials and Methods

Subjects. A total of 66 healthy subjects participated in the study. Three subjects were excluded due to technical errors during magnetic resonance image acquisition. Seven subjects did not meet the selection criterion that demanded a correct report of the outcome on at least 90% of all trials in the gambling task. The remaining 56 subjects (27 women; mean,
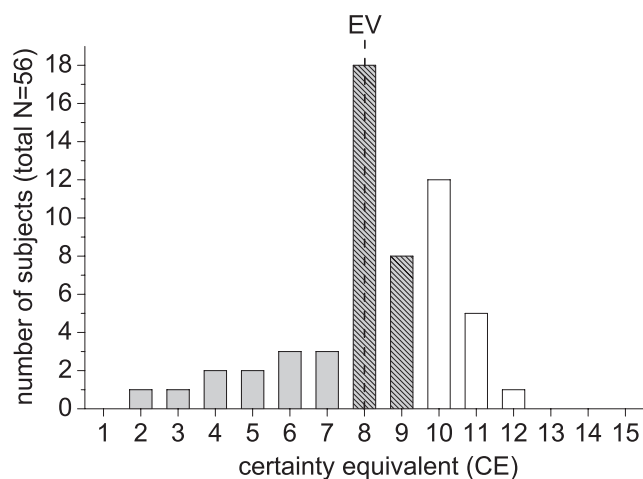
Figure 1. Distribution of CEs in the sample. The CE is the lowest safe payment a subject prefers over playing a lottery. The difference between the CE and the lottery's expected value (EV; indicated by the dashed line) is used to infer subjects' risk preferences. Columns are shaded in gray for risk averters and white for risk seekers. The diagonal line pattern indicates risk-neutral subjects who were indifferent around the EV.

25.4 ± 3.2 SD years) answered correctly on a mean of 96.3 ± 2.54% SD of all trials. They won on a mean 50 ± 4.64% SD of all trials.

Risk preference measure. To determine a subject's risk type (risk averse, risk neutral, risk seeking), we elicited certainty equivalents (CEs), i.e., the lowest sure amount of money a subject prefers over playing a lottery. For this, subjects were presented with 15 binary choices between a sure win (increasing from 1 to 15 points) and a 50% chance lottery of winning 16 points or nothing (Dohmen et al., 2011). Initially, all subjects preferred the lottery. The choice at which subjects switched to the sure win defines the CE. The CE is monotonic— but typically not linear—in risk preferences as measured by different economic models. As such, this measure allows us to identify risk averters, risk neutrals, and risk seekers in a model-independent manner. Figure 1 illustrates that the distribution of CE values in our sample is skewed in a way that single CE values above the lottery's expected value are overrepresented; in addition to the nonlinearity assumption regarding the CE, the skewed CE distribution further supports the use of the risk type as a between-subject factor in the subsequent analysis of the imaging data.

Basic economic theory defines subjects who prefer the safe option when faced with a fair gamble as risk averse and those who are indifferent, i.e., they are equally likely to choose the safe option or to gamble, as risk neutral. Applied to the iterative binary-choice lottery paradigm used in this study, risk averters switch to the safe option before the sure win reaches the lottery's expected value (CE, <8; Fig. 1). Subjects who are indifferent about the fair gamble, i.e., they either choose the safe option (CE, 8) or gamble and switch at the next choice (CE, 9), are defined as risk neutral. Subjects who choose to gamble beyond that (CE, >9) are defined as risk seeking. For the main analyses, data from risk-averse and risk-neutral subjects were combined ($n = 37$) and compared with risk-seeking subjects ($n = 19$). To see whether we could replicate the effects with only the extreme risk preference types, we excluded all risk-neutral subjects. For an exploratory analysis, we also compared all three types: 11 risk-averse, 26 risk-neutral, and 19 risk-seeking subjects.

Before answering the lottery questionnaire, subjects were informed that, in addition to their win from the gambling task, 1 of the 15 lotteries would be randomly chosen and paid off with 1 point representing €0.50. On average, participants won a mean €4 ± 3.51 SD in the lottery task.

fMRI gambling task. The event-related functional magnetic resonance imaging (fMRI) paradigm consisted of a simple card gambling task (Preuschoff et al., 2006) (Fig. 2). In this task, participants first place a bet on whether a second presented card will be higher or lower than the first card. Then they watch the two cards being consecutively drawn from a randomly shuffled deck of 10 nonface spade cards, ranging from ace to 10. With display of the first card, the expected reward is updated and

results in a reward prediction error. Similarly, the anticipation risk, i.e., the expected outcome variance, is updated and results in a risk prediction error. The updated anticipation risk maintains until the second card is shown and the gamble is resolved, which results in another risk prediction error. At the same time, a reward prediction error occurs if the expected reward deviates from the actual gamble outcome. Finally, participants are asked to indicate whether they have won or lost the gamble to control for attention. Participants who failed to report the correct outcome on at least 90% of all trials were excluded.

Within-trial sequences were randomized without replacement. Pseudorandomization of between-trial sequences ensured a relatively wide range of outcome probabilities and outcome variances as well as the exclusion of repetition trials. Each participant played three sessions with 30 trials each, always starting off with an initial balance of €25. One euro was at stake in each trial. To reduce wealth effects participants were informed that one of the three sessions would be randomly chosen at the end of the experiment and that the associated winning balance would be paid off. If payments came to <€10 per hour (including payoff from the lottery task; see above), the difference was compensated for.

MRI data acquisition. All MRI sessions were run on a Siemens Trio 3.0 T scanner with a standard eight-channel head coil. Structural scans included T1-weighted images (TR, 1570 ms; TE, 3.42 ms; flip angle, 15°; 1 mm slices). Three functional sessions were run, each of which started off with a localizer scan followed by the gambling paradigm implemented in Presentation (Neurobehavioral Systems; http://www.neurobs.com) during which T2*-weighted echoplanar images were collected (TR, 2500 ms; TE, 33 ms; flip angle, 90°; 33 2 mm slices in ascending order; field of view, 230 mm; voxel size, 1.8 × 1.8 × 2.0 mm; ~288 volumes per session).

fMRI data preprocessing. Preprocessing of the functional images was implemented in the MATLAB based (MathWorks) software Statistical Parametric Mapping 5 (SPM5; http://www.fil.ion.ucl.ac.uk/spm) and included realignment, coregistration, slice timing correction, and spatial smoothing with an 8 mm filter as well as normalization on the MNI standard (Evans et al., 1993) (http://www.bic.mni.mcgill.ca).

fMRI data analyses. The statistical analyses of the fMRI data in SPM8 were based on five general linear models (GLMs; 1–5). In short, GLM 1 uses parametric modulations of the risk regressors. GLMs 2 and 3 model anticipation risk and risk prediction error, respectively, as a five-level factor (one for each possible risk or error value). GLMs 4 and 5 model anticipation risk and risk prediction error, respectively, as a two-level factor based on categories of high and low risk, and positive and negative errors. All GLMs were estimated using a hemodynamic response function and a high-pass filter of 128 Hz as well as correction for autocorrelations and will be described in more detail below.

GLM 1 includes parametrically modulated regressors of interest for risk prediction error and anticipation risk and was set up to run random effects analyses at the group level to locate the mean neural sensitivities to each regressor of interest (cf. Preuschoff et al., 2006, 2008). For this GLM, we divided the onset regressor at Card 1 into a short epoch of 1 s and a long anticipatory epoch of ~6 s, depending on the jittered interval until the onset of Card 2. As a whole, the GLM includes the modulating parameters expected reward and risk prediction error at onset of Card 1, risk prediction during the long epoch in anticipation of Card 2 (anticipation risk), and risk prediction error and outcome (win or loss) at Card 2. As nuisance regressors we used visual and motor activation during bet and attention check, and individual movement regressors. Table 1 gives an overview of all mathematical parameters, Table 2 shows how the parameters of interest correspond to reward probability at Card 1. For more details, the reader is referred to Preuschoff et al. (2008). Single-subject contrasts were computed for the parametrically modulated regressor for risk prediction error at Card 1 and anticipation risk. These contrasts were then tested in separate random effects analyses (one-sample t tests) at the group level. We applied a voxelwise statistical threshold of $p < 0.05$ and corrected for multiple comparisons based on familywise error (FWE) control for the whole-brain volume. Based on previous findings, we expected to find activations in ventral striatum and anterior insula for anticipation risk (Preuschoff et al., 2006, 2008) (Table 3), and activations in anterior insula and inferior frontal gyrus for risk prediction error (Preuschoff et al., 2008) (see Table 5).
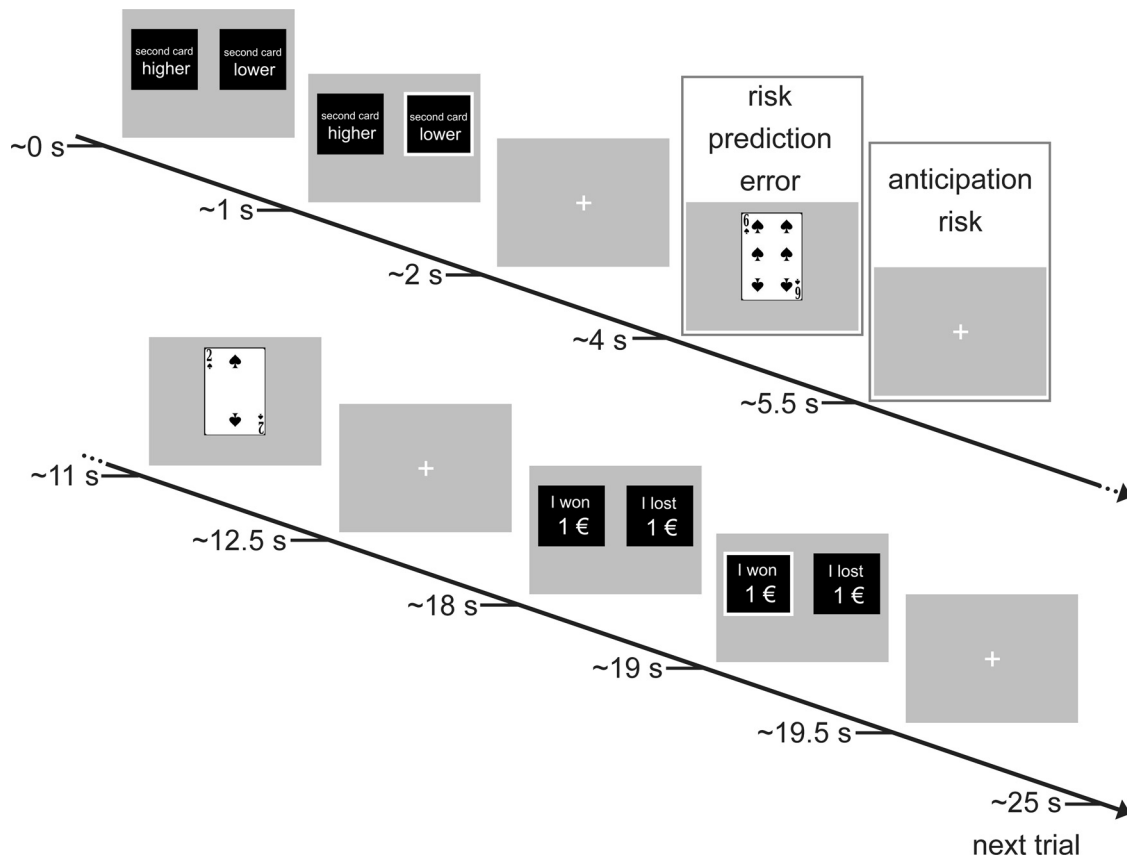
Figure 2. Time line of the gambling paradigm. Participants start off with an initial endowment of €25. In 30 trials they first place a bet of €1 on whether the second card will be higher or lower than the first card. After placing the bet, they watch the first card being drawn, followed 6 s later by the second card. Finally, they are asked to indicate whether they have won or lost the gamble. With display of the first card, risk is partially resolved, and a risk prediction error occurs. The anticipation risk is maintained until display of the second card, when the gamble is resolved.

Table 1. Formal definitions and numerical examples of reward and risk predictions assuming the subject bets "lower"

Term and definition	Example
Expected reward at Card 1 $P_i = E[P_2 = p * 1 + (1 - p) * (-1)]$	$5/9 * 1 + 4/9 * (-1) = 1/9$
Anticipation risk before Card 1 (constant) EP_1^2	$1/10 * [(-1)^2 + (-7/9)^2 + (-5/9)^2 + \ldots + 1] = 33/81$
Risk prediction error at Card 1 $P_1^2 - EP_1^2$	$(1/9)^2 - 33/81 = -32/81$
Anticipation risk before Card 2 $E(P_2 - P_1)^2$	$5/9 * (8/9)^2 + 4/9 * (-10/9)^2 = 80/81$
Outcome at Card 2 $P_2 \in \{1, -1\}$	1
Risk prediction error at Card 2 $(P_2 - P_1)^2 - E(P_2 - P_1)^2$	$(8/9)^2 - 80/81 = -16/81$

The first card is a 6, the second card is a 2 (Fig. 2). P denotes the reward probability based on the bet and the value of Card 1. P_i denotes the reward information revealed at card i. $E[x]$ denotes the expected value of a given parameter x.

Table 2. Overview of decision parameters after display of Card 1

Reward probability p at Card 1	Risk prediction error at Card 1	Anticipation risk before Card 2
1	0.59	0
0.89	0.20	0.40
0.78	−0.10	0.69
0.67	−0.30	0.89
0.56	−0.40	0.99
0.44	−0.40	0.99
0.33	−0.30	0.89
0.22	−0.10	0.69
0.11	0.20	0.40
0	0.59	0

The table shows how the 10 possible reward probabilities at Card 1 correspond to the parameter values for risk prediction error and anticipation risk, rounded to two decimal places.

To illustrate how the estimates of the blood oxygen level-dependent (BOLD) activation in our sample correspond to the anticipation risk values and the risk prediction error values, respectively, GLM 1 was modified in two ways: in GLM 2, the anticipatory epoch before Card 2 was modeled by five separate onset regressors for each anticipation risk level; in GLM 3 the onset of Card 1 was modeled by five separate onset regressors for each risk prediction error value (Table 2). Next, we extracted the mean β estimates for each regressor and each subject using a 10 mm sphere around the group peak voxels from the previous random

effects analyses (for coordinates, see Tables 3, 5). Finally, we averaged the β estimates across all subjects and plotted the overall mean estimates against the anticipation risk values (Fig. 3) and the risk prediction error values (see Fig. 5). Note that these β estimates were only used for descriptive plots.

To test whether the individual CE is reflected in the mean BOLD sensitivity toward anticipation risk and risk prediction error, we included the CE as a covariate in the above described one-sample t tests on the respective modulated regressors from GLM 1. For each covariate analysis, we applied a statistical threshold of $p < 0.05$, FWE corrected for a priori-defined regions of interest (Tables 3, 5). To avoid circular inferences, we defined these regions of interest based on activation peaks reported in previous studies: we used 10 mm spheres around coordinates of ventral striatum and anterior insula in tests of anticipation risk (Table 3) (Preuschoff et al., 2006, 2008) and coordinates of anterior insula and

Table 3. Regions positively correlated with anticipation risk (i.e., expected outcome variance)

Region	Laterality	MNI coordinates			Cluster size k_E	Max. t
		x	y	z		
Ventral striatum	L	−9	12	−3	145	10.00*
From Preuschoff et al. (2006)		(−10)	(−3)	(−4)		
	R	9	9	−3	112	9.21*
		12	−3	−4		
Anterior insula	L	−30	24	3	104	10.16*
From Preuschoff et al. (2006)		(−30)	(21)	(11)		
From Preuschoff et al. (2008)		(−31)	(22)	(10)		
	R	39	21	0	109	8.80*
From Preuschoff et al. (2006)		(31)	(24)	(11)		
From Preuschoff et al. (2008)		(33)	(21)	(10)		

Results from the random effects analysis are shown, and regions of interest for small-volume correction (with a sphere of 10 mm) as reported by Preuschoff et al. (2006, 2008) are shown in parentheses. Height threshold, $t = 5.28$; extent threshold, 100 voxels.

*$p < 0.05$ (activation survives whole-brain correction for multiple comparisons based on FWE control at the peak level).

inferior frontal gyrus in tests of risk prediction error (see Table 5) (Preuschoff et al., 2008).

In addition to the mean sensitivity as captured by the parametrically modulated regressors, we were particularly interested in how risk preferences are integrated in low- and high-risk conditions as well as in negative and positive risk prediction errors. Based on time course analyses on the current and previous data sets (cf. Preuschoff et al., 2008), we argue that these categories can be related to separate cognitive processes that evoke distinct BOLD signals and can thus be used as meaningful regressors in a GLM. To accommodate for the mixed within- and between-subject design in the same group-level analysis, we used factorial designs. As a between-subject factor, we used the subjects' risk preference type because the underlying CE values are assumed to be nonlinear with risk preferences and because the distribution of CE values in our sample is skewed (see above, Risk preference measure). The within-subject factors were modeled by categories of low and high anticipation risk, and negative and positive risk prediction error, respectively. Table 2 illustrates how the anticipation risk values correspond to reward probability p: the two lower anticipation risk values were merged to the category "low risk" (corresponding to $p \leq 0.3$, or $0.7 \leq p$), and the two upper anticipation risk values were merged to the category "high risk" (corresponding to $0.3 < p < 0.7$). Trials with zero risk were excluded because they represent reward information independent from risk. Based on these categories, we generated GLM 4 for each subject, which included a regressor for low- and high-risk trials, respectively, again including all of the other above-mentioned regressors of no interest. Single-subject contrasts were computed for low and high anticipation risk and entered into a group-level analysis. On second level, we ran a two-by-two ANOVA with a between-subject factor risk preference (averse and seeking) and a within-subject factor anticipation risk (low and high). The ANOVA was designed as a flexible factorial model with an internal replication factor for the subjects. We first combined the data from risk-averse and risk-neutral subjects ($n = 36$) and compared them with risk-seeking subjects ($n = 19$), and then replicated the effect with only the extreme risk preference types (11 risk averters and 19 risk seekers) after excluding all risk-neutral subjects (Fig. 1). In both cases we tested the interaction and applied a threshold of $p < 0.05$, FWE corrected for the a priori-defined regions of interest (Table 3).

The same procedure was repeated to test how risk preferences are reflected in negative and positive risk prediction errors. The risk prediction error is negative when the current risk is smaller than expected and positive when the current risk is larger than expected. We generated GLM 5, which included a regressor for negative risk prediction errors (corresponding to $p \leq 0.2$, or $0.8 \leq p$) and positive risk prediction errors at Card 1 (corresponding to $0.2 < p < 0.8$), as well as the above mentioned regressors of no interest Table 2 illustrates how the risk prediction error values correspond to reward probability. Similar to the group-level analysis on GLM 4, we ran a two-by-two ANOVA with between-subject

factor risk preference (averse and seeking) and within-subject factor risk prediction error (negative and positive). The analysis was done for the combined dataset from risk-averse and risk-neutral subjects and then replicated without risk-neutral subjects. Again, we tested for an interaction and applied a threshold of $p < 0.05$, FWE corrected for the a priori-defined regions of interest (see Table 5). To show the nature of the interaction effects from both factorial analyses, we then extracted the β estimates as follows. For each subject, we identified the peak voxel with the maximum Z-score within the above-described regions of interest in the single-subject contrast images for low and high anticipation risk separately, as estimated in GLM 4 (the same for negative and positive risk prediction errors as modeled in GLM 5). Next, we extracted the β value that corresponded to the specific peak voxel. Then, we averaged the individual β estimates across risk-averse and risk-seeking subjects separately and plotted the respective mean estimates against low and high risk (Fig. 4) and negative and positive risk prediction errors (Fig. 6). These plots aim to illustrate the nature of the two-by-two interaction effects supported by the above described inferential statistics. Note that the individual peak voxel selection was based on the separate contrasts of low and high anticipation risk (negative and positive risk prediction errors) and is not biased by the interaction effect shown on group level or by a main effect of high risk being greater than low risk.

To see whether we could find a similar interaction effect of the risk preference types across all five levels of anticipation risk and risk prediction error, we ran two three-by-five factorial analyses based on GLM 2 and GLM 3. Because these analyses divide the original parametrically modulated regressors from GLM 1 into five separate onset regressors that essentially model the same cognitive process, the regressors are not perfectly independent and their power to explain meaningful variance in the BOLD signal decreases significantly. Also, some of the risk values were less likely to occur in the experimental paradigm than others. As a consequence, the subject groups are not equally represented in each regressor, especially for the extreme risk prediction error values. We were thus concerned that the five onset regressors in GLMs 2 and 3 would underestimate the BOLD signal across the increasing risk values, and we therefore refer to these analyses as exploratory. The first exploratory analysis used the single-subject contrasts for five levels of anticipation risk as modeled in GLM 2. At the group level, we ran a three-by-five ANOVA with a between-subject factor risk preference (averse, neutral, seeking) and a within-subject factor anticipation risk (0, 0.4, 0.69, 0.89, 0.99). The same procedure was repeated in the second exploratory analysis in which we used the single-subject contrasts for five levels of risk prediction error as modeled in GLM 3. We ran a three-by-five ANOVA with a between-subject factor risk preference (averse, neutral, seeking) and a within-subject factor risk prediction error (−0.4, −0.3, −0.1, 0.2, 0.59). In both cases we tested for a linear parametric interaction and applied FWE correction to the regions of interest at $p < 0.05$ (Tables 3, 5). For descriptive purposes, we extracted the individual mean β estimates as described for GLM 2 and GLM 3 above, averaged them for each risk type, and plotted them together with the overall average in Figures 3 and 5.

Risk-learning simulation. To obtain a qualitative picture of the impact that asymmetrically represented risk prediction errors have on prospective risk perception, we implemented a simple reinforcement learning algorithm for risk. The algorithm is based on the assumption that risk learning parallels reward learning in that risk perception is updated with every risk prediction error. We simulated 1000 consecutive trials for two groups of subjects and two levels of risk: a low-risk condition with a small expected outcome variance and a high-risk condition with a high expected outcome variance. In the subject group "risk seekers" (1), the current risk estimate is correctly increased or decreased after positive or negative risk prediction errors, respectively. In the subject group "risk averters" (2), any negative risk prediction errors are underweighted by 20%, and thus the subsequent risk estimate is maladjusted. An illustration of this risk-learning simulation can be seen in the inset of Figure 4.

Statistical analyses of behavioral data. The reaction times during the fMRI gambling task and the exact stimuli sequence of each trial were logged by the Presentation software and then analyzed in MATLAB. Two-sample t tests were run to compare (1) the sum of switch trials with that of stay trials and (2) the reaction times after losing with those after
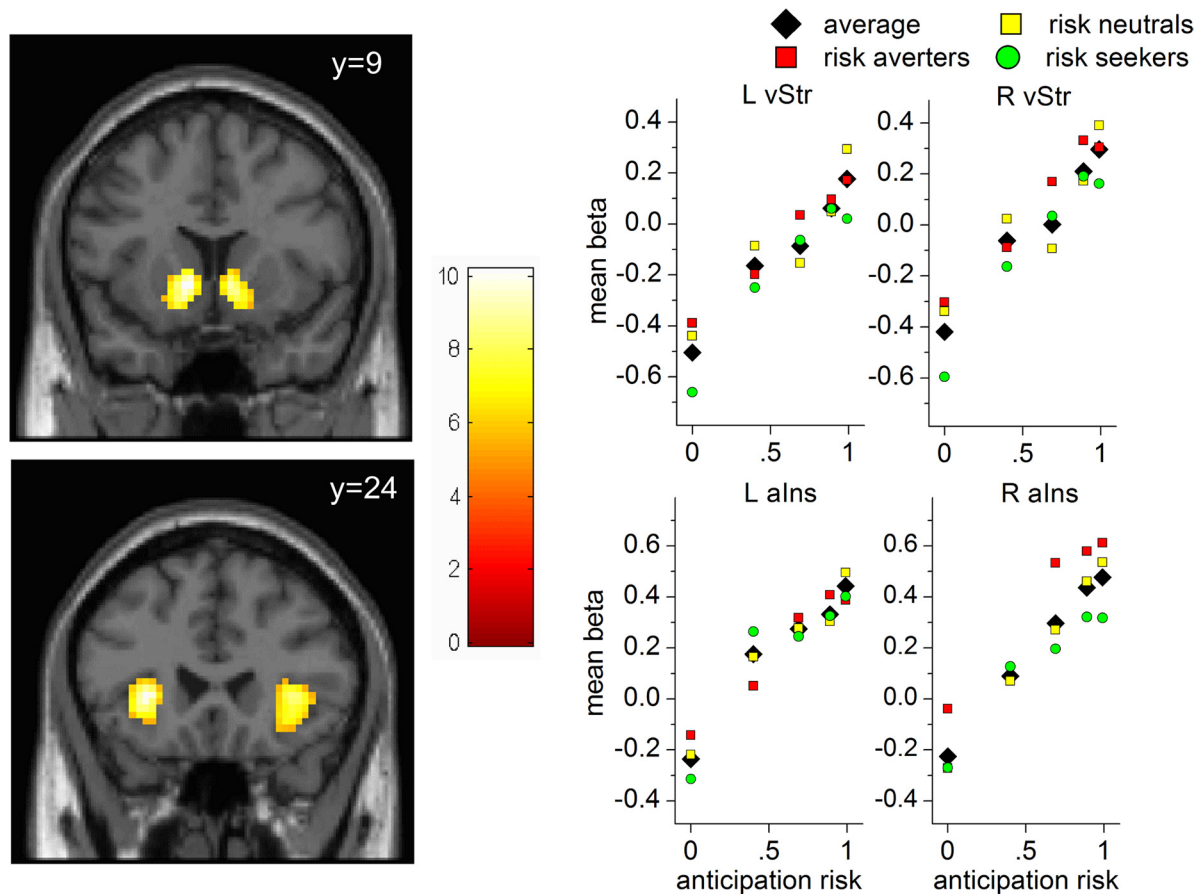
Figure 3. Anticipation risk coding. Neural activation in bilateral ventral striatum (vStr; top) and anterior insula (aIns; bottom) during anticipation of the second card that correlates with anticipation risk, i.e., expected outcome variance. The figures (left) show statistical parametric maps of the random effects analysis, color coded for the *t* values as indicated by the color bar ($p < 0.05$, FWE corrected for whole-brain volume). The graphs (right) show mean β estimates in the respective structures plotted against five anticipation risk values, averaged over all subjects and for each risk preference type individually. On average, neural responses are lowest for sure wins and losses and highest for trials with uncertain outcomes. Risk-averse and risk-neutral subjects show greater responses to higher risk values than risk-seeking subjects.

winning a gamble. Paired *t* tests were used to test whether the sum of switch trials differed with respect to the preceding gamble outcome. Data from risk-averse and risk-neutral subjects were combined for this analysis and compared with risk-seeking subjects.

Results

Neuroimaging results

Increased anticipation risk signal reflects risk aversion
First, to identify the regions that process risk during the anticipation of the outcome, we tested the associated single-subject contrasts estimated in GLM 1 in a one-sample *t* test at the group level. We confirmed that bilateral ventral striatum as well as anterior insula code for anticipation risk ($p < 0.05$, FWE corrected for whole-brain volume; Fig. 3; Table 3). The right panel in Figure 3 was derived from GLM 2 for descriptive purposes; it illustrates that the mean activation in these regions increases linearly with anticipation risk.

Second, to test whether the individual CE is reflected in the mean neural sensitivity to anticipation risk in the regions of interest (Table 3), we added the CE as a covariate to the one-sample *t* test. We applied FWE correction at $p < 0.05$ to the regions of interest and found no correlation of the CE with the mean sensitivity to anticipation risk.

Third, we were particularly interested in the interaction between neural sensitivities to low and high anticipation risk and the individual risk preference, i.e., whether a subject is risk

averse or risk seeking. Based on time course analyses reported elsewhere (Preuschoff et al., 2008), we argue that low- and high-risk trials can be related to independent cognitive processes that evoke distinct BOLD signals, and can thus be modeled as separate regressors in a GLM. Contrast images for low and high anticipation risk were estimated in GLM 4, and their interaction with the risk preference type (risk averters and risk neutrals compared to risk seekers) was tested in a mixed within- and between-subject two-by-two ANOVA. The a priori-defined target regions that are reported in Table 3 (Preuschoff et al., 2006, 2008) reveal a significant interaction effect ($p < 0.05$, FWE corrected for small volumes; Table 4). Figure 4 illustrates the nature of this two-by-two interaction: when anticipating high-risk gambles, risk averters show a stronger BOLD response in ventral striatum and anterior insula compared to risk seekers. This interaction effect remained significant in all regions of interest after we excluded risk-neutral subjects ($p < 0.05$, FWE corrected for small volumes). Moreover, the regions of interest showed a main effect of risk aversion (risk averters > risk seekers; $p < 0.05$, FWE corrected for small volumes).

To confirm this interaction effect across all five levels of anticipation risk, we tested the contrast images estimated in GLM 2 for a linear parametric interaction with all three risk preference types in a three-by-five ANOVA. Because we were concerned that the

regressors in GLM 2 were not independent and would underestimate the linear effect of anticipation risk in the BOLD signal, we refer to this analysis as exploratory. The results support our previous finding: all regions of interest showed a significant interaction effect ($p < 0.05$, FWE corrected for small volumes; Table 4). Figure 3 illustrates that risk averters and risk neutrals show a stronger increase of the BOLD signal with increasing risk levels compared to risk seekers.

Risk averters show an elevated risk prediction error signal

Next, we tested for the BOLD effect of risk prediction error at Card 1 in the associated contrast images estimated in GLM 1 in a one-sample t test over all subjects. We found BOLD activation in bilateral anterior insula, bilateral IFG, the anterior cingulate cortex (ACC), and bilateral ventral striatum ($p < 0.05$, FWE corrected for whole-brain volume; Fig. 5; Table 5). The plot in the right panel of Figure 5 is based on contrast images from GLM 3; it was generated for descriptive purposes and illustrates that mean activation estimates in the indicated brain structures increase linearly with risk prediction error.

We then included the individual CE as a covariate in the one-sample t test and applied a threshold of $p < 0.05$, FWE corrected for the regions of interest (Table 5). We found a correlation of the CE with the mean sensitivity to risk prediction errors in a single voxel in the left ventral striatum. Because this effect is most likely driven by the skewed CE distribution in our sample, which causes an overrepresentation of higher CE values, we refrain from interpreting this result.

Finally, we were interested in the interaction between the neural sensitivities to negative and positive risk prediction errors and the individual risk preferences. Time course analyses in previous studies (cf. Preuschoff et al., 2008) suggest that these categories evoke distinct BOLD signals and can be used as separate regressors. The corresponding contrast images estimated in GLM 5 were entered into a two-by-two ANOVA to test for an interaction of the risk preference type (risk averters and risk neutrals compared to risk seekers) and risk prediction errors (negative and positive). All a priori-defined search volumes that are reported in Table 5 (Preuschoff et al., 2008), i.e., bilateral anterior insula and right IFG, showed a significant interaction effect ($p < 0.05$, FWE corrected for small volumes; Table 6). In a test on whole-brain activation, the ACC showed the same interaction effect ($p < 0.05$, FWE corrected for whole-brain volume). The plots of peak β activations in Figure 6 help us to understand the nature of this interaction: risk seekers respond more to positive than to negative risk prediction errors. Risk averters, by contrast, respond with an elevated, but equally strong BOLD signal to both negative and positive risk prediction errors. After we excluded the subjects who were indifferent around the lottery's expected value (Fig. 1), we found the same interaction effect in right IFG and ventral striatum at a threshold of $p < 0.005$ (uncorrected). At the same thres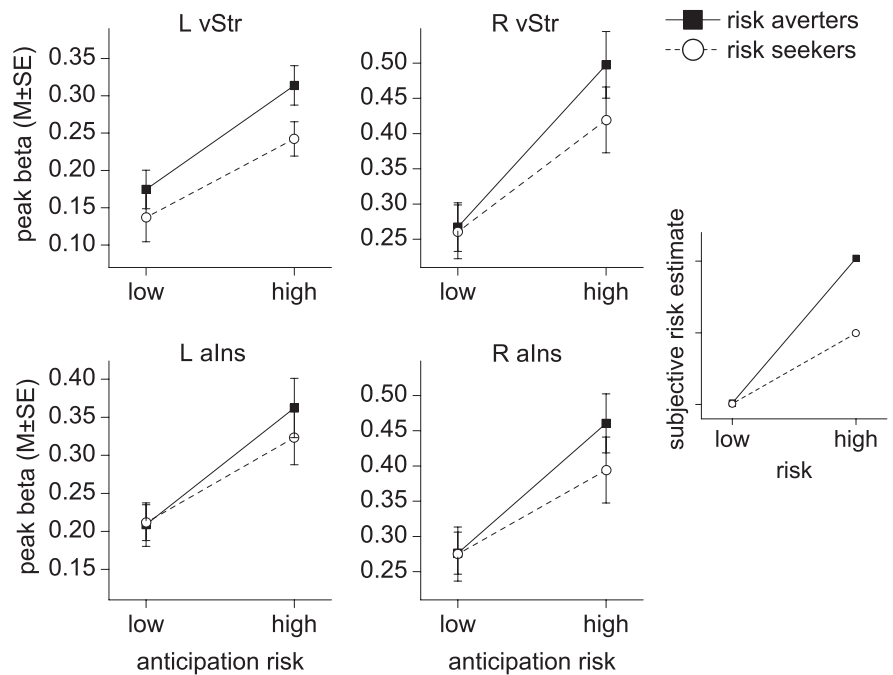hold, the bilateral anterior insula, right IFG, ACC, and ventral striatum also showed a main effect of risk aversion (averters > seekers).

In an additional analysis, we wanted to confirm the interaction of risk preference types across all five levels of risk prediction error in a three-by-five ANOVA. Because the regressors in GLM 3 are not perfectly independent and the model might underestimate the linear BOLD effect across the risk prediction error values, we refer to this analysis as exploratory. In line with our previous results, we find a linear parametric interaction effect in the IFG (FWE corrected at $p < 0.05$ for a priori-defined search volumes; Table 5), in the right anterior insula (FWE corrected at $p < 0.05$ for the whole-brain volume), and in the left anterior insula and the ACC at a threshold of $p < 0.001$ (uncorrected; Table 6). The plots in Figure 5 show that risk-averse subjects have a generally elevated BOLD signal compared to risk-seeking subjects, whereas risk seekers show a stronger increase across the risk prediction error values.

Risk-learning simulation

To understand how an asymmetric representation of positive and negative risk prediction errors would influence the evaluation of prospective risk, we simulated a simple risk-learning model. The risk-seekers algorithm (1) incorporated the correct adjustment of the current risk estimate with every encountered risk prediction error. The risk-averters algorithm (2) included a nonsufficient decrease of the current risk estimate with every negative risk prediction error (Fig. 6). The inset in Figure 4 illustrates that the impact of asymmetric risk prediction error representations (Algorithm 2) on risk prediction changes with the experienced risk level: a disproportionate representation of negative risk prediction errors leads to an

Figure 4. Interaction effect of risk preferences and anticipation risk coding in a two-by-two ANOVA. During high-risk trials, the neural response to anticipation risk is stronger in risk averters than in risk seekers (interaction effect, $p < 0.05$, FWE corrected for small volumes). Plots show the mean (M) and SE of the individual peak β estimates within the indicated search volume. Data from risk-averse and risk-neutral subjects were combined for these plots. The inset (right) illustrates subjective risk estimates based on a simple reinforcement learning model using asymmetric risk prediction errors. This model predicts the interaction effect that we observe for anticipation risk and risk preferences. vStr, Ventral striatum; aIns, anterior insula.
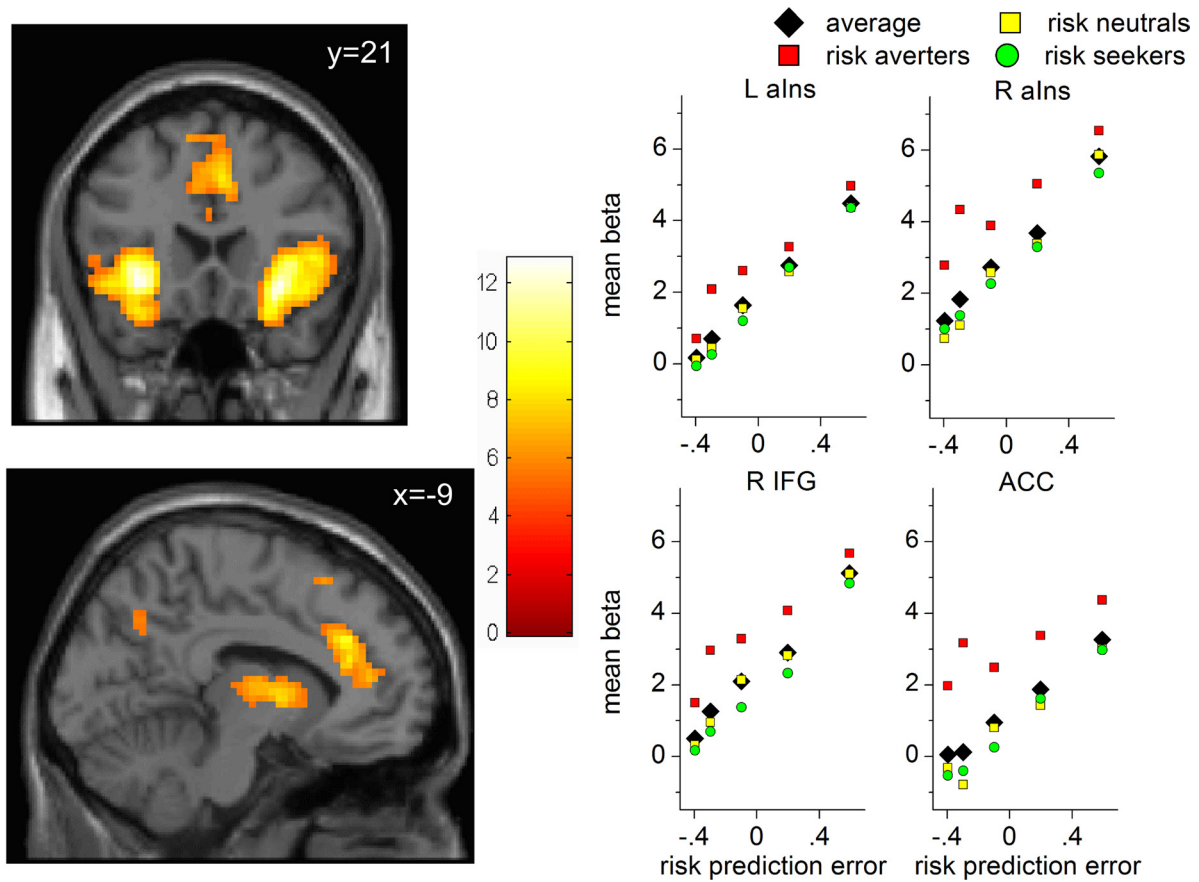
Figure 5. Risk prediction error coding. Neural activation in bilateral anterior insula (aIns; top), ACC, and vStr (bottom) correlate positively with deviations from previous risk predictions, i.e., risk prediction error. The figures (left) show statistical parametric maps of the random effects analysis, color coded for the t values as indicated by the color bar ($p < 0.05$, FWE corrected for whole-brain volume). The graphs (right) show mean β estimates in the respective structures plotted against five risk prediction error values, averaged over all subjects and for each risk preference type individually. On average, neural responses increase linearly with risk prediction error. Risk-averse subjects respond more strongly than risk-seeking subjects to any risk prediction error.

Table 4. Regions positively correlated with the interaction of risk preferences and anticipation risk

Region	Laterality	MNI coordinates			Cluster size k_E	Max stat t
		x	y	Z		
Ventral striatum	L	−12	6	−6	16	4.32*
		(−9)	(6)	(0)	(9)	(4.28)*
	R	12	6	0	12	3.31*
		(9)	(6)	(−3)	(5)	(3.39)*
Anterior insula	L	−30	24	9	58	3.95*
		(−24)	(24)	(6)	(8)	(3.43)*
	R	33	24	9	56	4.44*
		(36)	(24)	(9)	(12)	(3.31)*

Results from the two-by-two ANOVA are shown. Data from risk-averse and risk-neutral subjects were combined for this analysis. Results from the exploratory three-by-five analysis are shown in parentheses. Inclusion threshold, $t = 2.62$ (2.59).

*$p < 0.05$. The activation survives small-volume correction for multiple comparisons based on FWE control at the peak level for a priori-defined regions of interest as reported by Preuschoff et al. (2006, 2008) (Table 3).

Table 5. Regions positively correlated with risk prediction error

Region	Laterality	MNI coordinates			Cluster size k_E	Max stat t
		x	y	z		
Anterior insula	L	−33	21	−9	445	12.82*
From Preuschoff et al. (2008)		(−31)	(15)	(−2)		
Superior temporal gyrus		−27	15	−21		8.34*
Inferior frontal gyrus		−48	15	3		7.20*
Anterior insula	R	33	24	−6	632	12.50*
From Preuschoff et al. (2008)		(32)	(16)	(−3)		
Superior temporal gyrus		48	21	3		8.88*
Inferior frontal gyrus		51	36	−9		7.28*
From Preuschoff et al. (2008)		(48)	(17)	(18)		
Anterior cingulate cortex	L	−6	33	27	709	9.97*
	R	9	36	24		8.49*
Ventral striatum, caudate	L	−9	0	3	209	7.79*
	R	6	−3	3		7.06*

Results from the random effects analysis are shown, and regions of interest for small-volume correction (with a sphere of 10 mm) as reported by Preuschoff et al. (2008) are shown in parentheses. Height threshold, $t = 5.27$; extent threshold, 100 voxels.

*$p < 0.05$ (activation survives whole-brain correction for multiple comparisons based on FWE control at the peak level).

even greater overestimation of prospective risk under high-risk conditions than under low-risk conditions.

Behavioral results

We also analyzed subjects' behavior at the time of placing the bet in the gambling task. Notably, the task was designed so that strategic behavior would not improve the potential outcome (Preuschoff et al., 2006). The total number of switch compared to stay trials, i.e., how often the subjects change their bet as opposed to repeatedly choosing the same bet, did not differ significantly

($t_{(55)} = -1.328$; $p = 0.19$). Overall reaction times when placing the bet were independent from the previous gamble outcome ($t_{(55)} = 0.712$; $p = 0.48$). When we investigated the risk preference groups separately (combining data from risk averters and risk neutrals in line with the main fMRI analyses), though, we found that risk-averse subjects are unaffected by preceding outcomes, whereas risk seekers change their bet significantly more

Table 6. Regions positively correlated with the interaction of individual risk preferences and two categories of risk prediction error (negative and positive)

Region	Laterality	MNI coordinates			Cluster size k_E	Max stat t
		x	y	z		
Anterior insula	L	−33	21	−9	51	5.35*
		(−42)	(24)	(−3)	(77)	(4.35)**
	R	30	21	−9	57	4.53*
		(48)	(21)	(6)	(3)	(4.82)***
Inferior frontal gyrus	R	51	21	12	19	3.45*
		(51)	(21)	(12)	(9)	(4.20)*
Anterior cingulate cortex	R	6	42	15	3	5.24***
		(6)	(42)	(12)	(16)	(4.23)**

Results from the two-by-two ANOVA are shown. Data from risk-averse and risk-neutral subjects were combined for this analysis. Results from the exploratory three-by-five analysis are shown in parentheses.

*$p < 0.05$. The activation survives small-volume correction for multiple comparisons based on FWE control at the peak level for a priori-defined regions of interest as reported by Preuschoff et al. (2008) (Table 5). Inclusion threshold, $t = 2.62$ (2.59).

**$p < 0.001$. The activation does not survive correction for multiple comparisons. For the sake of completeness the activation in this region of interest is reported at $p < 0.001$, uncorrected (height threshold $t = 3.12$).

***$p < 0.05$. The activation survives whole-brain correction for multiple comparisons based on FWE control at the peak level. Height threshold, $t = 4.92$ (4.76).
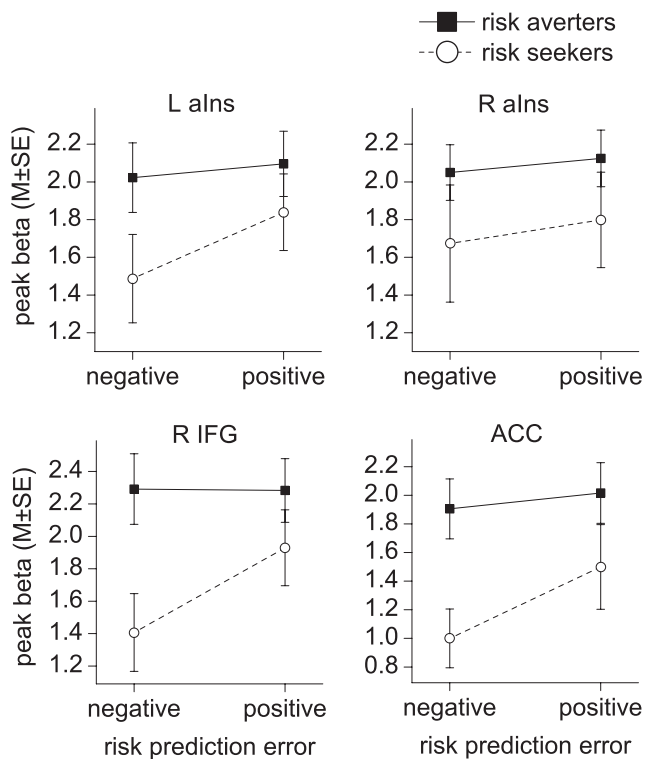

Figure 6. Interaction effect of risk preferences and risk prediction error coding in a two-by-two ANOVA. Risk averters show an elevated but equally strong BOLD signal to both positive and negative risk prediction errors. Risk seekers respond significantly less to gambles that are not as risky as expected, reflected by negative risk prediction errors (interaction effect, $p < 0.05$, FWE corrected for small volumes; ACC, $p < 0.05$, FWE corrected for whole-brain volume). Plots show the mean (M) and SE of the individual peak β estimates within the indicated search volume. Data from risk-averse and risk-neutral subjects were combined for these plots. aIns, Anterior insula; L, left; R, right.

often after losing than after winning a gamble ($t_{(18)} = −6.263$; $p = 0.006$) and are slower when switching their bet ($t_{(18)} = −2.606$; $p = 0.018$).

Discussion

We examined the effects of individual risk preferences on the neural coding of risk with a combination of an fMRI gambling

paradigm and a behavioral lottery task. We found that risk aversion covaries with an increased anticipation risk signal in the ventral striatum and anterior insula. Moreover, risk aversion is reflected in an elevated response to negative risk prediction errors in anterior insula, IFG, and ACC. Both risk evaluation signals occur independent of choice, thus implying automatic cognitive processes.

The ventral striatum has been reported to code for expected reward and the corresponding reward prediction error on the one hand (Schultz et al., 1997; Knutson et al., 2001; Preuschoff et al., 2006) and anticipation risk on the other (Fiorillo et al., 2003; Preuschoff et al., 2006). It is important to note that anticipation risk, i.e., the expected outcome variance, also plays a significant role in reward learning because it represents a predictability measure (Preuschoff and Bossaerts, 2007). In other words, in an uncertain, highly variable environment, a deviation from the forecast is not surprising. Under low risk conditions, though, where the outcome can be predicted with high confidence, any prediction error is much more informative and should thus have more impact. On a neural level, the sustained anticipation risk signal in the ventral striatum may serve as such an amplifier for the subsequent reward prediction error.

The anterior insula, in contrast, is involved in risk learning (Preuschoff et al., 2008). Our study shows that the anterior insula reflects both anticipation risk and the corresponding risk prediction error. This is an important function because only under uncertain conditions (e.g., risk) can an organism experience prediction errors and thus learning (Rescorla and Wagner, 1972). Hence, both risk-seeking and risk-averse behavior can be appropriate at times. The choice of which risk behavior to engage in is likely to depend on the individual risk evaluation. Our results show that the individual risk behavior indeed corresponds to differential evaluation and updating of risk information. In risk averters, the response to negative risk prediction errors is elevated, implying that negative errors are not perceived as strongly (i.e., as negatively) as in risk seekers. Interestingly, this can be reconciled with the finding that risk averters show a stronger response in high-risk situations. Assume that risk learning parallels reward learning in that risk is updated based on past risk prediction errors. A decision maker who places too little emphasis on negative as compared to positive risk prediction errors will on average overestimate risk because she or he does not sufficiently reduce her or his estimate after experiencing negative errors. This effect is stronger for high-risk than for low-risk situations as seen in the inset of Figure 4, which qualitatively resembles our results. As such, a simple reinforcement learning model using asymmetric risk prediction errors predicts the interaction effect that we observe for anticipation risk and risk preferences in anterior insula.

In addition, the anterior insula has been associated with the representation and updating of internal body states (Paulus and Stein, 2006) as well as the anticipation of aversive stimuli (Yáguez et al., 2005). An altered interoceptive prediction mechanism in the anterior insula has even been discussed to account for anxiety (Paulus and Stein, 2006): the avoidance of fearful stimuli that is characteristic for anxiety patients may be caused by the exaggerated anticipation of feared stimuli. This increased anticipatory anxiety itself may be a consequence of an enhanced prediction error feedback that leads to an overestimation of prospective fearful occurrences. The purpose of avoidance behaviors may be to attenuate the experience of prediction errors. Likewise, risk aversion may be due to a hypersensitive anticipation risk signal that in turn is driven by a continuous risk prediction error signal.

Accordingly, our data show that risk averters respond equally strongly to positive and negative risk prediction errors, and have an increased BOLD signal for the anticipation of risky stimuli. Risk-averse behavior may thus serve the purpose to avoid future risk prediction errors, coded by the anterior insula, and the accompanying subjective feeling (Craig, 2002, 2009).

At first sight, the purely mathematical tracking of objective risk (Preuschoff et al., 2008) and the generation of interoceptive prediction signals (Craig, 2002, 2005) may seem contradictory (Bossaerts, 2010). But a wide range of literature shows that emotional reward anticipations are necessary for advantageous decision making (Bechara et al., 1994, 1997). Given its role in the generation of subjective feelings (Craig, 2009), the anterior insula may code for both objective and subjective risk, i.e., mathematical risk prediction and individual risk preference, respectively (Singer et al., 2009). Other imaging studies that included choice during risk anticipation did not find effects of risk preferences in the insular cortex (Christopoulos et al., 2009). Here, we focus on passive risk evaluation and do observe an altered insula signal in risk averters. This suggests that insula signals of risk learning, rather than signals during choice, may indeed reflect, at least in part, subjective risk.

It is of interest whether this subjective risk signal can predict risk behavior. A previous study implied such a link between anticipatory insula activation and subsequent risk behavior (Xue et al., 2010): depending on whether subjects had previously chosen to gamble and won, the evaluation of the subsequent fair gamble elicited stronger insula activation. The same enhanced evaluation signal predicted risky choices on the following trial. Since trials in our current design were independent from each other, we refrain from interpreting between-trial effects on BOLD responses. However, the paradigm could easily be adapted to include risk-taking behavior after an experimental manipulation of anticipation risk and risk prediction error.

In a previous study, Tobler et al. (2007) found risk-related activation in IFG that covaries with individual risk preferences. The authors reported an area in the IFG that not only codes for expected reward, but also shows increased activation with risk only for risk-seeking subjects. Given its role in reward and uncertainty coding (Tobler et al., 2007) and in the updating of risk information (Preuschoff et al., 2008), the IFG seems to be an important node in the risk processing network that is biased by the individual risk attitude. This is supported by our result of a differential IFG signal for risk prediction errors in risk seekers compared to risk averters.

The ACC is assumed to be involved in processes of conflict monitoring and error detection (Michelet et al., 2009). It has been speculated that the ACC activation provides a general alertness for error detection and supports the avoidance of future errors. The differential ACC signal for risk prediction errors that we detect for risk seekers may thus represent a reduced monitoring of negative compared to positive risk prediction errors. This could bias risk seekers' learning experience toward positive errors that contribute to their general risk-seeking strategy. In contrast, risk averters show an elevated ACC signal that does not discriminate appropriately between positive and negative risk prediction errors. This may indicate an increased level of alertness for both risk prediction errors.

In summary, this study shows that risk preferences are not only reflected in behavioral differences but already in the underlying evaluation process. More specifically, risk aversion is associated with increased risk anticipation as well as enhanced neurophysiological signals for any risk prediction errors in risky

situations without an actual choice. Thus, risk-averse subjects seem to be more sensitive to the variance of outcomes, which implies that they perceive risk more strongly. This result complements previously reported effects of risk preferences on risk processes during or after choices. The striatal anticipation risk signal may modulate the impact of subsequent reward learning, whereas the anterior insula activation may represent the integration of objective and subjective risk and the corresponding anticipation of somatosensory states. We suggest that the exaggerated risk prediction error signals in risk averters drive the overestimation of prospective risk. This is supported by the risk averters' increased risk anticipation signal. Risk preferences may thus be due to alterations on different steps of risk processing, such as the emotional risk evaluation as well as the cognitive inhibition of risk-taking behavior.

Our study supports the notion that individual risk attitudes should be taken into account for an ecologically valid risk-learning model. As stated in the introduction, the finding that risk averters and risk seekers differ in their risk evaluation has practical implications for many domains: the interindividual heterogeneity in risk learning could, for example, inform the prediction of stock market investments. The result that risk seekers show a weaker response to risk might also motivate changes in health campaigns; focusing on health risks in smoking prevention campaigns might be ineffective in risk-seeking smokers. In general, public policies dealing with decision making under risk should encompass how the subjective risk evaluation might bias the actual decision.

References

Bechara A, Damasio AR, Damasio H, Anderson SW (1994) Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50:7–15. CrossRef Medline

Bechara A, Damasio H, Tranel D, Damasio AR (1997) Deciding advantageously before knowing the advantageous strategy. Science 275:1293–1295. CrossRef Medline

Blais AR, Weber EU (2006) A domain-specific risk-taking (DOSPERT) scale for adult populations. Judgment Decision Making 1:33–47.

Bossaerts P (2010) Risk and risk prediction error signals in anterior insula. Brain Struct Funct 214:645–653. CrossRef Medline

Christopoulos GI, Tobler PN, Bossaerts P, Dolan RJ, Schultz W (2009) Neural correlates of value, risk, and risk aversion contributing to decision making under risk. J Neurosci 29:12574–12583. CrossRef Medline

Craig AD (2002) How do you feel? Interoception: the sense of the physiological condition of the body. Nat Rev Neurosci 3:655–666. Medline

Craig AD (2005) Forebrain emotional asymmetry: a neuroanatomical basis? Trends Cogn Sci 9:566–571. CrossRef Medline

Craig AD (2009) How do you feel—now? The anterior insula and human awareness. Nat Rev Neurosci 10:59–70. CrossRef Medline

Dohmen T, Falk A, Huffman D, Sunde U, Schupp J, Wagner GG (2011) Individual risk attitudes: measurement, determinants, and behavioral consequences. J Eur Econom Assoc 9:522–550. CrossRef

Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, Peters TM (1993) 3D statistical neuroanatomical models from 305 MRI volumes. Proc IEEE-Nuclear Science Symposium and Medical Imaging Conference 3:1813–1817.

Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. Science 299:1898–1902. CrossRef Medline

Gianotti LR, Knoch D, Faber PL, Lehmann D, Pascual-Marqui RD, Diezi C, Schoch C, Eisenegger C, Fehr E (2009) Tonic activity level in the right prefrontal cortex predicts individuals' risk taking. Psychol Sci 20:33–38. CrossRef Medline

Guiso L, Paiella M (2005) The role of risk aversion in predicting individual behaviour. Report, Economic Research Department, Bank of Italy, Rome.

Knutson B, Fong GW, Adams CM, Varner JL, Hommer D (2001) Dissociation of reward anticipation and outcome with event-related fMRI. Neuroreport 12:3683–3687. CrossRef Medline

Markowitz HM (1952) Portfolio selection. J Finance 7:77–91. CrossRef

Michelet T, Bioulac B, Guehl D, Goillandeau MB, Burbaud P (2009) Single medial prefrontal neurons cope with error. PLoS One 4:e6240. CrossRef Medline

Mohr PN, Biele G, Heekeren HR (2010) Neural processing of risk. J Neurosci 30:6613–6619. CrossRef Medline

Paulus MP, Stein MB (2006) An insular view of anxiety. Biol Psychiatry 60:383–387. CrossRef Medline

Preuschoff K, Bossaerts P (2007) Adding prediction risk to the theory of reward learning. Ann N Y Acad Sci 1104:135–146. CrossRef Medline

Preuschoff K, Bossaerts P, Quartz SR (2006) Neural differentiation of expected reward and risk in human subcortical structures. Neuron 51:381–390. CrossRef Medline

Preuschoff K, Quartz SR, Bossaerts P (2008) Human insula activation reflects risk prediction errors as well as risk. J Neurosci 28:2745–2752. CrossRef Medline

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcemen. In: Classical conditioning II: current research and theory (Black AH, Prokasy WF, eds), pp 64–99. East Norwalk, CT: Appleton Century Crofts.

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593–1599. CrossRef Medline

Singer T, Critchley HD, Preuschoff K (2009) A common role of insula in feelings, empathy and uncertainty. Trends Cogn Sci 13:334–340. CrossRef Medline

Tobler PN, O'Doherty JP, Dolan RJ, Schultz W (2007) Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. J Neurophysiol 97:1621–1632. Medline

Tversky A, Kahneman D (1992) Advances in prospect theory: Cumulative representation of uncertainty. J Risk Uncertainty 5:297–323. CrossRef

Weber EU, Hsee C (1998) Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes towards perceived risk. Management Sci 44:1205–1217. CrossRef

Xue G, Lu Z, Levin IP, Bechara A (2010) The impact of prior risk experiences on subsequent risky decision-making: the role of the insula. Neuroimage 50:709–716. CrossRef Medline

Yáguez L, Coen S, Gregory LJ, Amaro E Jr, Altman C, Brammer MJ, Bullmore ET, Williams SC, Aziz Q (2005) Brain response to visceral aversive conditioning: a functional magnetic resonance imaging study. Gastroenterology 128:1819–1829. CrossRef Medline