Neural correlates of risk prediction error during reinforcement learning in humans

Mathieu d'Acremont [a,*], Zhong-Lin Lu [b], Xiangrui Li [b], Martial Van der Linden [a], Antoine Bechara [b]

[a] National Centre of Competence in Research (NCCR) in Affective Sciences, University of Geneva, Rue des Battoirs 7, CH-1205 Geneva, Switzerland
[b] Dana and David Dornsife Cognitive Neuroscience Imaging Center, Brain and Creativity Institute, Department of Psychology, University of Southern California, Los Angeles, CA 90089-1061, USA

ARTICLE INFO

ABSTRACT

Behavioral studies have shown for decades that humans are sensitive to risk when making decisions. More recently, brain activities have been shown to be correlated with risky choices. But an important gap needs to be filled: How does the human brain learn which decisions are risky? In cognitive neuroscience, reinforcement learning has never been used to estimate reward variance, a common measure of risk in economics and psychology. It is thus unknown which brain regions are involved in risk learning. To address this question, participants completed a decision-making task during fMRI. They chose repetitively from four decks of cards and each selection was followed by a stochastic payoff. Expected reward and risk differed among the decks. Participants' aim was to maximize payoffs. Risk and reward prediction errors were calculated after each payoff based on a novel reinforcement learning model. For reward prediction error, the strongest correlation was found with the BOLD response in the striatum. For risk prediction error, the strongest correlation was found with the BOLD responses in the insula and inferior frontal gyrus. We conclude that risk and reward prediction errors are processed by distinct neural circuits during reinforcement learning. Additional analyses revealed that the BOLD response in the inferior frontal gyrus was more pronounced for risk aversive participants, suggesting that this region also serves to inhibit risky choices.

© 2009 Elsevier Inc. All rights reserved.

The simplest model of decision-making posits that options with higher expected rewards are preferred over options with lower expected rewards. But much evidence from experimental psychology and economics indicate that humans and animals are sensitive to risks in addition to expected rewards when making decisions. A usual way to measure risk in finance is to define it as a function of outcome variability (Rothschild and Stiglitz, 1970). For instance, in the mean–variance approach, the value of a portfolio is traded off against its return variance (Markowitz, 1952). Variance is also central to modern portfolio theory like Capital Asset Pricing Model (CAPM, Sharpe, 1964) and to the modelling of financial time series like Autoregressive Conditional Heteroskedasticity (ARCH, Engle, 1982).

In behavioral science, risk measures derived from variance, as the coefficient of variation, have been found to explain the decision-making of humans and animals (Weber et al., 2004; McCoy and Platt, 2005). In addition, there is ample evidence that the central nervous system reacts differentially to risk. Single cell recording studies in monkeys have highlighted a subsample of neurons that increases their activities after the presentation of conditioned stimuli only when these stimuli predict variable outcomes (Fiorillo et al., 2003). Functional magnetic resonance imaging (fMRI) studies in humans

have revealed neural correlates of risky stimuli or decisions in various brain regions including the striatum, the insula, the inferior frontal gyrus, the lateral orbitofrontal cortex, and the anterior cingulate cortex (Paulus et al., 2003; Huettel et al., 2005; Kuhnen and Knutson, 2005; Rolls et al., 2007; Tobler et al., 2007).

Ambiguity refers to situations in which probabilities of potential outcome of a decision are incomplete or unknown. Under ambiguity, decision-making has been correlated to activities in the amygdala, the orbitofrontal cortex, the inferior frontal gyrus, and the insula (Huettel et al., 2006; Hsu et al., 2005). It should be noted that ambiguity carries risk because outcomes are uncertain (Rode et al., 1999) and this may explain the partial overlap of brain activation for decision-making under ambiguity and risk.

There is cumulative evidence that the brain responds to stimuli that are associated with risky outcomes. But it remains unclear how the brain learns the association between risk and certain stimulus or action. Reward prediction error is central to many reinforcement algorithms (Sutton and Barto, 1998). This error is the difference between the predicted and the experienced reward. The predicted reward is updated after each trial based on the reward prediction error and a learning rate. After sufficient number of trials and with an appropriate learning rate, the predicted reward converges to the expected value of the reward. We already know that reward prediction errors are represented in the central nervous system. Indeed, single unit recording in monkey has revealed that activities of the midbrain dopaminergic neurons follow the time course of reward

* Corresponding author. École Polytechnique Fédérale de Lausanne, Laboratory for Decision-Making under Uncertainty, Odyssea, Station 5, CH-1015 Lausanne, Switzerland. Fax: +4121 693 00 20.
E-mail address: Mathieu.Dacremont@epfl.ch (M. d'Acremont).

prediction error during classical conditioning (Schultz et al., 1997). Activity in the human striatum has been related to reward prediction error during classical and instrumental conditioning (O'Doherty et al., 2004; McClure et al., 2003).

Recently a new reinforcement learning model has been developed to estimate the outcome variance (Preushoff and Bossaerts, 2007). In this algorithm a second prediction error is computed: the risk prediction error. With this signal, the agent is able to estimate risk. The risk prediction error is defined as the difference between the predicted and the realized risk. The realized risk is the squared reward prediction error. The predicted risk is updated after each trial based on the risk prediction error and a learning rate. After a sufficient number of trials and with a correct learning rate, the predicted risk converges to the expected value of the risk to be realized, which equals to the outcome variance.

One fMRI study has explored the neural correlates of risk prediction error (Preuschoff et al., 2008). In this study, a card was drawn from a deck of 10 (numbered from 1 to 10) followed several seconds later by a second card (drawing without replacement). Participants bet one dollar on whether the second card would be lower than the first card. Before any card is drawn, the expected reward is 0 and the variance is 1. But after the first card has been drawn, expected value and variance need to be updated. For instance, if the first card is a 9, the expected value increases and the variance decreases (because it is likely that the second card will be below 9). So there is a positive reward prediction error and a negative risk prediction error. When the second card is drawn, the expected value and variance need to be updated again, which produces a new reward and risk prediction error. The main result showed that risk prediction error was related to activation in the insula.

It should be noted that in Preuschoff et al. (2008) study, reward probability was explicit. It can be computed at any time step based on the cards drawn (decision-making under uncertainty). In addition, participants passively observed the card selected by the computer. The question arises if the brain also computes risk prediction error when probabilities are unknown and expected value and variance need to be learned through experience (decision-making under ambiguity). Secondly, we can wonder if risk prediction errors are observed when participants are free to select among several options and if this signal serve to orient action policy. As such, the aim of this study was to localize risk learning signals in the human brain in an ambiguous situation offering free choices. To do so, we used a reinforcement learning algorithm that can estimate expected risk. Such an algorithm has never been used in brain imaging studies.

Based on previous results (O'Doherty et al., 2004; McClure et al., 2003), we hypothesized that reward prediction error would be related to activity in the ventral striatum. Based on the study of Preuschoff et al. (2008), we hypothesized that activity in the insula would be related to risk prediction error. Because participants were free to select the option they preferred, they did not only need to evaluate risk, but also to adapt their action policy as a function of risk (e.g., avoid risky choices in risk aversive individuals). Thus we may also highlight brain regions that control the action policy toward risk. To explore these research questions, participants completed four versions of the Iowa Gambling Task during fMRI, in both active and control conditions.

One advantage of selecting the Iowa Gambling Task is that it has been used in numerous neuropsychological studies and it appears to have ecological validity: participants who make suboptimal decisions in the Iowa Gambling Task do the same in real life situations. Studies have shown that performance in the Iowa Gambling Task is impaired following lesions encompassing but not restricted to the medial orbitofrontal cortex (Bechara et al., 1997). However, only few studies have examined brain activities during the Iowa Gambling task (Fukui et al., 2005; Oya et al., 2005; Ernst et al., 2002). None has dissociated the expected value and variance of the decks and applied a reinforcement risk-sensitive learning algorithm. This is in striking contrast with the fact that the majority of neurological or psychological disorders for which a deficit in the Iowa Gambling Task has been observed are characterized by risk taking in real life situations (e.g., prefrontal lesions, Bechara et al., 1997 or antisocial behavior, van Honk et al., 2002; Anderson et al., 1999).

Materials and methods

Subjects

Eight students from the University of Southern California, 4 males and 4 females, all right-handed, with a mean age of 23.00 years (SD = 3.30 years) and normal or corrected-to-normal vision, participated in the study after MRI safety screening and full informed consent. The experimental procedures were approved by the Institutional Review Board at the University of Southern California. Informed consent was obtained from all subjects.

Behavioural task

Four versions of the Iowa Gambling Task (Bechara et al., 1994) were used in this study: ABCD, KLMN, EFGH, and IJOP. Each version was played in two conditions: An ambiguous and a control condition. In the ABCD version and the ambiguous condition, the subject saw 4 decks of cards on a computer screen labeled A, B, C, and D (Fig. 1, left). There were 60 cards in each deck. In a deck, the back of all the cards were the same, but half of the cards have a red face, half a black face. Subjects used 4 buttons on an MRI-compatible response box to select one of the four decks. After each selection, the card was immediately turned face up and a gain or a gain and a loss is displayed with the following message: "Win X" or "Win X but loses Y". The feedback was displayed immediately after the selection and till the beginning of the next trial. The sum of the gain and loss after a card selection is referred to as the payoff. A green bar near the top of the screen displayed the cumulative monetary payoff. Each trial lasted 4 s. In the control condition, payoffs were displayed on each card so participants knew the outcome in advance (Fig. 1, right). Blocks of the ambiguous condition were interrupted with blocks of the control condition (Fig. 1, bottom). Each block was made of 20 trials and there was 5 blocks per condition in each fMRI run. Whereas we have tried in every possible way to make the stimuli and design identical to the original Iowa Gambling Task, we imposed a 4-second time limit on each trial in order to facilitate fMRI data collection: If the subject did not make any response 3.5 s after the card presentation, the computer program made a random response and proceed to give feedback and then start the next trial. Our subject missed only 0.2% of the trials in the ambiguous condition and 0.4% in the control condition.

The four versions of the IGT differ in the composition of the decks. In the ABCD version, the mean payoff is negative for decks A and B and positive for decks C and D. The payoff variability is greater for decks A and B compared to C and D. Thus an important feature of this version is that the bad decks in term of expected value, are also risky in term of variance (Fig. 2, ABCD). The summed payoff decrease linearly over series of 10 cards in decks A and B, thus they become more disadvantageous over time. The summed payoff increases linearly in decks C and D, thus they become more advantageous over time. In other words, the expected value is non-stationary and need to be updated. In the ABCD version, the selection of a card is always followed by a gain, sometimes also by a loss. The structure of the KLMN version is similar to that of the ABCD version, except that the position of the good decks is different and that the variance of the payoffs is higher for all decks. As a consequence, the expected values of the decks are more difficult to estimate in the KLMN version. To favor learning, the KLMN version was played right after the ABCD version.
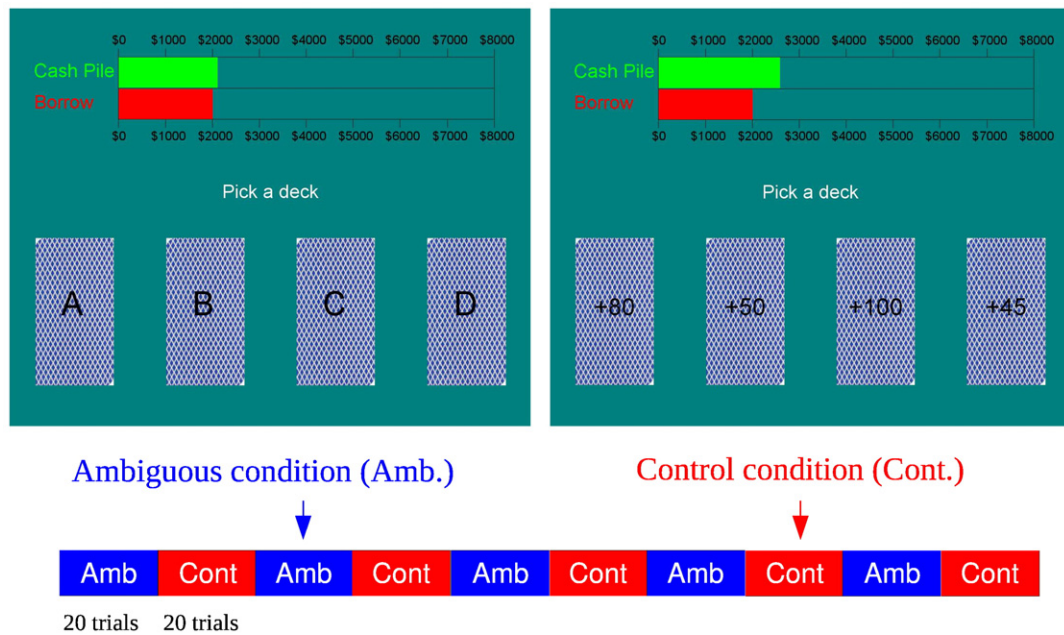
Fig. 1. Ambiguous (left) and control conditions (right) of the Iowa Gambling Task.

In the EFGH version, the mean payoff is negative for decks F and H and positive for decks E and G. The payoff variability is greater for decks E and G compared to F and H. Thus an important feature of this version is that the bad decks in term of expected value, are also safe in term of variance (Fig. 2, EFGH). The summed payoff decreases linearly over series of 10 cards in decks F and H, thus they become more disadvantageous over time. The summed payoff increases linearly in decks E and G, thus they become more advantageous over time. Thus here again, the expected value is non-stationary. In the EFGH version, the selection of a card is always followed by a loss, sometimes also by a gain. The structure of the IJOP version is similar to that of the EFGH version, except that the variance of payoff is higher for all decks. As a consequence, the expected values of the decks are more difficult to estimate in the IJOP version. The IJOP version was played right after the EFGH version.

Prior to the scanning session, participants were given instructions on the Iowa Gambling Task. It was explained that two decks were "worse" and that they should stay away from these decks in order to win. It was also explained that the computer would not change the deck composition after the game started (for details of the instructions, see Bechara et al., 2000). Participants were informed that another kind of task would run in between the primary task. For this control condition, they were instructed to select the card with the highest payoff. Participants ran 20 trials in a practice task similar to the IGT before the scanning session.

Image acquisition

MRI recording was performed using a standard birdcage head coil on a Siemens 3 T MAGNETON Trio MRI system housed in the Dana and David Dornsife Cognitive Neuroscience Imaging Center at University of Southern California. Subjects lay supine on the scanner bed, and viewed the back-projected visual displays through a built-in mirror on the head coil. Foam pads were used to minimize head motion. For each subject, sagittal images ($256 \times 256 \times 192$ voxels) of 1 mm^3 isotropic spatial resolution were obtained with a T1-weighted 3D MPRAGE
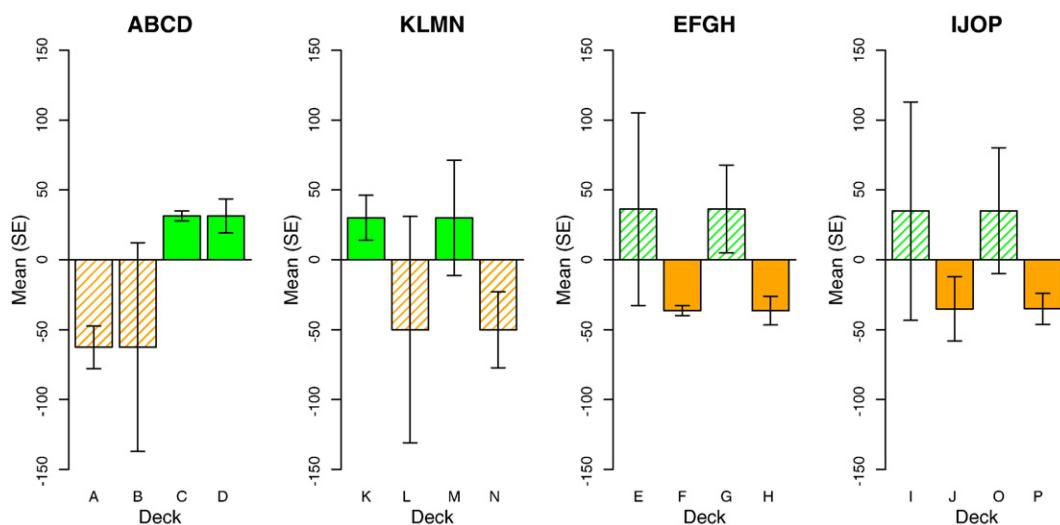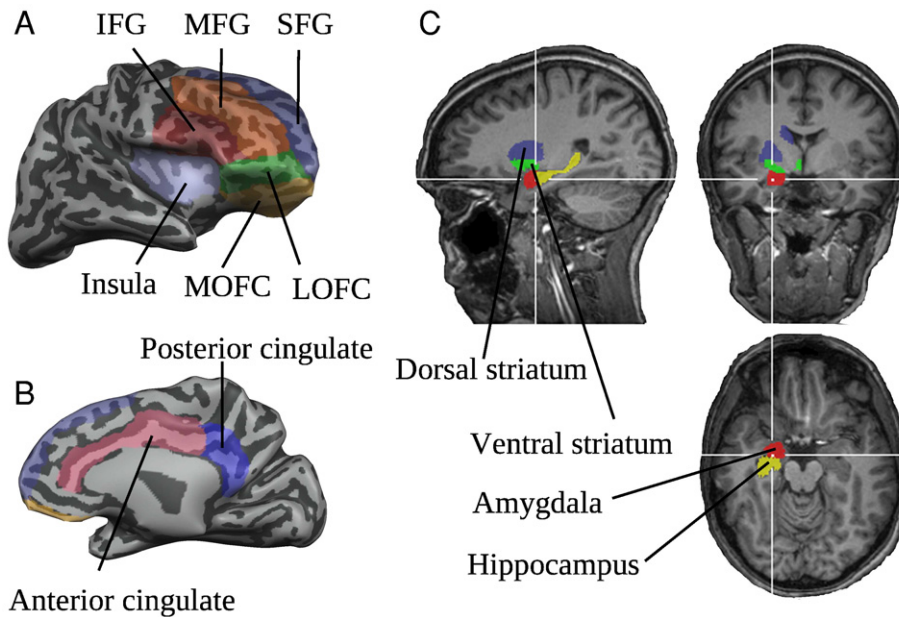

Fig. 2. Mean payoff and standard error (SE) by deck for the four versions of the Iowa Gambling Task. Each deck consisted of 60 cards. Good decks are in green and bad decks in orange. Safe decks are plain and risky decks are dashed.

Fig. 3. Regions of interest: SFG, Superior frontal gyrus; MFG, Middle frontal gyrus; IFG, Inferior frontal gyrus; LOFC, Lateral orbitofrontal cortex; MOFC, Medial orbitofrontal cortex.

sequence (TI = 900 ms, TR = 2070 ms, TE = 4.13 ms, flip angle = 7°). Blood-oxygenation-level-dependent (BOLD) responses were measured with a T2*-weighted echo-planar imaging (EPI) sequence (TR = 2000 ms, TE = 25 ms, flip angle = 90°, FOV = 192 × 192 mm, in-plane resolution = 64 × 64 pixels or 33 mm). Thirty-five interlaced coronal slices with a 3.5 mm (no gap) slice thickness were acquired.

Two task orders were used and counterbalanced between subjects: ABCD, KLMN, EFGH, IJOP and EFGH, IJOP, ABCD, KLMN. Four hundred eight volumes were recorded for each version of the task. The 4 first and last volumes of each run were recorded when participants fixed a cross in the center of the display. The presentation of the cards in the beginning of each trial was synchronized with the TR. Two volumes were acquired in each trial. Each subject participated in four 13.6-minute functional runs. Acquisition of structural images took place between the second and third functional runs (192 volumes). Each session lasted about 1.5 h.

Image analysis

All MRI- and fMRI-related data analyses were performed using BrainVoyager QX 1.10.4 (Brain Innovation, Maastricht, The Netherlands). The anatomical data for each subject were corrected for image intensity inhomogeneity, and transformed into the Talairach space (Talairach and Tournoux, 1988). The gray–white matter boundaries that resulted from gray–white matter segmentation were used to create a 3D surface model of the brain, which was then inflated to display both sulci and gyri on smooth surfaces of the two hemispheres. Twelve ROI in each hemisphere were selected in accordance with brain studies on decision-making under uncertainty (Knutson and Bossaerts, 2007). ROI were hand drawn by an expert based on anatomical features (Damasio, 2005). Regions on the cortical surface were first defined on the 3D inflated cortex (Figs. 3A and B) and cross validated with features on the 2D anatomical images. Regions not on the cortical surface, such as amygdala, striatum etc, were drawn directly on the 2D anatomical images (Fig. 3C). The average activation of all the voxels in each region of interest was used in ROI analyses.

Functional data were first pre-processed to correct for slice timing and head movement, followed by high-pass temporal filtering with a cutoff frequency of 3 cylces/run. The functional images were aligned to the structural images in the same session and constructed into a 4D volume in the Talairach space. For regions above the sinus, such as orbitofrontal region, there was some signal loss in functional images. This happened to four regions, the left Lateral OFC, right Lateral OFC, left Medial OFC, and right Medial OFC, with signal loss of about 4, 14, 31 and 30%, respectively. We adopted the two-gamma model in BrainVoyager for hemodynamic response function (Friston et al., 1999).

Reinforcement learning algorithm

A reinforcement learning model was used to model decision-making in the ambiguous condition.[1] The model includes prediction errors for both reward and risk, following the recent proposition of Preushoff and Bossaerts (2007). Predicted reward and risk were updated following a Rescorla–Wagner rule. Let the reward r_t denote the payoff in trial t. Let the value v_t be the predicted reward. Formally, we can write:

$$v_t = E(r_t), \tag{1}$$

where E is the expected value.

Let h_t denote the predicted risk at trial t. The predicted risk is defined as the variance of the payoff:

$$h_t = \mathrm{Var}(r_t). \tag{2}$$

Through learning, v_t is estimated by $\hat{v}_t$ and h_t by $\hat{h}_t$. Let δ_t denote the reward Prediction Error (PE) at trial t. This reward prediction error is the difference between the reward r_t and the estimated predicted reward $\hat{v}_t$:

$$\delta_t = r_t - \hat{v}_t. \tag{3}$$

Empirical data and theoretical considerations suggest that reward prediction error is scaled by predicted risk in the brain (Preushoff and Bossaerts, 2007; Tobler et al., 2005), allowing the learning rate to be

---

[1] For modelling, we ignored the time gaps created by the control blocks, which might have caused participants to forget what they had learned. However, the learning curve in the ambiguous condition and the model fit presented below suggest that the control blocks did not prevent reinforcement learning.

independent of the payoff variance. Let $\bar{\delta}_t$ denote scaled reward prediction error:

$$\bar{\delta}_t = \frac{\delta_t}{\sqrt{\hat{h}_t}}. \tag{4}$$

The Rescorla–Wagner rule is used to update $\hat{v}_t$ for the next trial:

$$\hat{v}_{t+1} = \hat{v}_t + \bar{k} \cdot \bar{\delta}_t, \tag{5}$$

where $\bar{k}$ is the reward learning rate.

After sufficient number of trials and with the appropriate learning rate, $\hat{v}_t$ converges to the expected value of r:

$$E(\hat{v}_t) \rightarrow E(r_t) = v_t. \tag{6}$$

Let ξ_t denote the risk prediction error at trial t. ξ_t is the difference between the squared reward prediction error and the estimated predicted risk:

$$\xi_t = \delta_t^2 - \hat{h}_t. \tag{7}$$

The Rescorla–Wagner rule is used to update $\hat{h}_t$ for the next trial:

$$\hat{h}_{t+1} = \hat{h}_t + k_{risk} \cdot \xi_t, \tag{8}$$

where k_{risk} is the risk learning rate. For the parsimony of the model, the same learning rate was used for reward and risk prediction error ($k = \bar{k} = k_{risk}$).

After sufficient number of trials and with the appropriate learning rate, $\hat{h}_t$ converges to the expected value of δ_t^2 and thus the variance of r:

$$E(\hat{h}_t) \rightarrow E(\delta_t^2) = Var(\delta_t) = Var(r_t). \tag{9}$$

Let $\hat{u}_{it}$ denote the utility of an option i at trial t, defined as a function of the predicted reward and risk associated with each option i:

$$\hat{u}_{it} = \hat{v}_{it} + l \cdot \sqrt{\hat{h}_{it}}, \tag{10}$$

where l is the risk preference (Bell, 1995).

Let π_{it} be the probability to select the option i in trial t. π_{it} was calculated with a softmax rule:

$$\pi_{it} = \frac{e^{\hat{u}_{it}}}{\sum_{j=1}^{n} e^{\hat{u}_{jt}}} \tag{11}$$

The reinforcement learning algorithm was implemented in C++ with the GNU Scientific Library (Galassi et al., 2006):

```
1: for game = 1:4 do //go through the 4 versions of the Iowa Gambling
      Task
2:    Initialize the deck selection probability πi = 0.25, i = 1,...,4
3:    Initialize estimated predicted reward v̂i = 0, i = 1,...,4
4:    Initialize estimated predicted risk ĥi = 1, i = 1,...,4
5:    for t = 1: n do //go through 100 trials of the game
6:       With a uniform random variable, select deck i according to πi,
         i = 1, ...,4
7:       Get the first outcome ra of the card in deck i
8:       Get the second outcome rb of the card in deck i
9:       Calculate the payoff, r = ra + rb
10:      Calculate the reward PE, δ = r − v̂i
11:      if first card in the deck and k>0
12:         Use the best estimate of the predicted risk, ĥi = δ²
13:      end if
14:      Scale the reward PE, δ̄ = δ / √ĥi
15:      Calculate the risk PE, ξ = δ² − ĥi
16:      Update the predicted reward, v̂i = v̂i + k · δ̄
17:      Update the predicted risk, ĥi = ĥi + k · ξ
18:      if it is the last card in the deck then
19:         Make the selection of deck i impossible v̂i = − ∞
20:      end if
21:      Calculate utility, ûi = v̂i + l · √ĥi
22:      Calculate deck selection probabilities, πi = e^ûi / ∑⁴ⱼ₌₁ e^ûj,
         i = 1, ...,4
23:   end for
24: end for
```

The reinforcement learning algorithm was fitted to the 400 decisions made in the ambiguous condition. The learning rate k and the risk preference l were estimated separately for each subject by maximizing the loglikelihood:

$$MLL = \max_{k,l} \sum_{t=1}^{400} \log \pi_{t,i_t}. \tag{12}$$

with i_t, the deck chosen at trial t, $t = 1,...,400$, $i_t \in \{1,2,3,4\}$; and π_{t,i_t} probability for selecting deck i_t in trial t.

It appears that the loglikelihood is not globally concave. In this case, the use of a direct search method like the Nelder–Mead algorithm can lead to local maxima and is inappropriate. Therefore, a local search method called Threshold Accepting was used (Winker and Gilli, 2004). The principle of this heuristic is to start the search with a random parameter combination and to move repetitively (n steps) to another combination (neighbor). The neighbor is accepted as the new combination if it has 1) a higher likelihood or 2) a lower likelihood and the likelihood difference is smaller than a given threshold. This search is repeated for decreasing threshold (n thresholds). The last neighbor found for the last threshold is retained as the solution. The solution is searched several times (n restarts) and the best solution is retained. The global maximum was search with 2000 steps, repeated for 10 thresholds, and for 30 restarts. Parameter constraints were $k \in [0,1]$ and $l \in [-0.01,0.01]$. The heuristic was implemented in C++.

Results

Reinforcement learning

We first check if there was evidence of learning in the ambiguous condition. The probability to select good decks was computed for each

Good Deck Selection
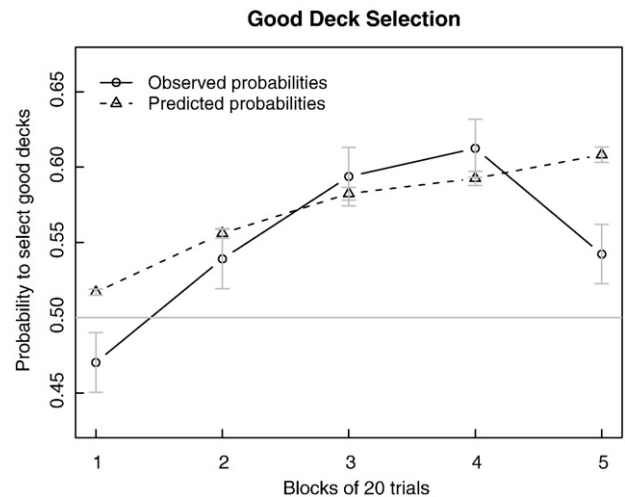
Fig. 4. Probability of selecting good decks in the ambiguous condition as a function of blocks, average across the four versions of the task. Circles are the probabilities computed from participant's choices, triangles are the probabilities computed from the reinforcement learning model.

Table 1
Reinforcement learning model estimation.

Subject	k	l	MLL	Lower	Upper
1	1.00e−01	−1.84e−03	−504.64	−521.28	−482.80
2	4.46e−03	3.32e−03	−546.12	−550.56	−535.29
3	3.59e−05	−1.64e−03	−553.05	−554.52	−548.31
4	1.91e−03	3.57e−03	−546.14	−551.36	−537.13
5	2.65e−02	1.22e−03	−543.19	−549.66	−532.50
6	1.20e−01	−1.52e−03	−466.12	−502.16	−456.46
7	4.31e−03	2.76e−03	−550.88	−552.79	−539.39
8	2.81e−02	7.00e−04	−545.29	−550.91	−536.70

k is the learning rate and l the risk preference parameter. *MLL* is the maximum likelihood estimated from subject's decisions. *Upper* and *Lower* define the 90% Confidence Interval of the simulated *MLL* (bootstrap).

Table 2
Comparison with a No-Risk model.

| Subject	MLL	MLL No-Risk	χ^2	p
1	−504.64*	−510.89	12.49	0.00
2	−546.12*	−554.52	16.80	0.00
3	−553.05	−554.52	2.94	0.09
4	−546.14*	−551.92	11.56	0.00
5	−543.19*	−545.49	4.60	0.03
6	−466.12*	−470.25	8.26	0.00
7	−550.88*	−554.52	7.28	0.01
8	−545.29	−545.97	1.35	0.25

*The model including risk is significantly better at $\alpha = 0.05$. *MLL No-Risk* is the maximum likelihood of the model with no risk (l fixed to 0).

block of 20 trials across all four versions of the task. Results showed that the probability to select the good decks increased over time, with however a decline of performance in the last block (Fig. 4, circles). A generalized mixed linear model for Bernoulli distribution was estimated with the glmmPQL function of R (R Development Core Team, 2007). Subject was entered as a random factor and trial (1 to 100) as a fixed regressor. The effect of trial on the probability to select good decks was significant, $t(3191) = 3.41$, $p < 0.001$, confirming the learning effect.

Model parameters were estimated with Threshold Accepting for each subject and are reported in Table 1. The mean learning rate was significantly greater than 0, suggesting that participants updated their estimate of the expected value and variance of the payoff over the course of each version of the task, $M = 0.036$, $SD = 0.047$, $t(7) = 2.13$, $p = 0.04$ (one-sided test).[2] The risk preference parameter was not significantly different from 0, $M = 0.00082$, $SD = 0.00228$, $t(7) = 1.02$, $p = 0.34$ (two-sided test). This shows that subjects overall were risk neutral. However the relatively big standard deviation (compared to the mean) also shows that there were important individual differences in risk preference. The probability to select the good decks predicted by the estimated reinforcement learning model is plotted in Fig. 4. There are differences between the observed and the predicted probabilities, but the model reproduces the general learning trend.

Parametric bootstrap was used to check the model fit: (1) fix the estimated parameters in the reinforcement learning algorithm, (2) generate 1000 simulations of the 400 decisions (100 decisions/ version × 4 versions of IGT) for each subject, (3) estimate parameters of the reinforcement learning model from the 1000 simulations. This simulation gives the maximum loglikelihood (MLL) distribution for decisions that follows the theoretical model. If the subject MLL lies within the 90% Confidence Interval of the simulated MLL, this suggests that the theoretical model is a good approximation of human decision-making in the Iowa Gambling Task. Based on this criterion, the theoretical model was accepted for all subjects (Table 1).

To test whether participants took risk into account in making their decisions, a simplified model was estimated after fixing the value of l to 0. In this simplified model, risk is ignored; decisions are made solely based on the expected values. The improvement of the goodness of fit from the model without risk to the model with risk was significant for 6 of the 8 subjects according to a χ^2 test, $\chi^2 = -2*(\text{MLL}_{\text{nested}} - \text{MLL})$ (Table 2). These results indicate that not all, but a majority of subjects included risk in their action policy.

---

[2] The mean learning rate appears to be relatively small. It should be noted that the learning rate cannot be directly compared with previous studies because we used a novel reinforcement algorithm. In particular, the reward prediction error was scaled by the reward SD. However, Oya et al. (2005) have applied a reinforcement learning algorithm (with no risk estimation) to account for the behavior of a neurosurgical patient who showed normal performance in the ABCD version of the Iowa Gambling Task. They found a learning rate of 0.076. This is above our mean learning rate but within our SD range.
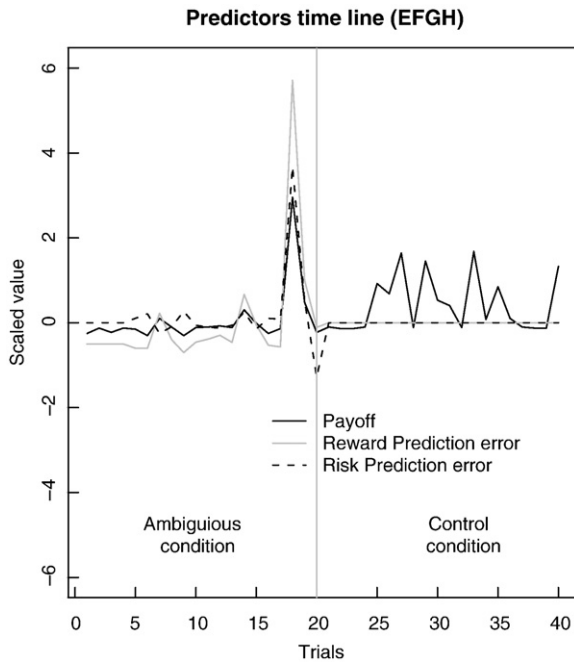
Risk and reward prediction error

For brain activation analysis, risk prediction error was transformed to share the same unit as reward prediction error by taking the signed square root of the absolute value, $\text{sign}(\xi) \cdot \sqrt{|\xi|}$. Reward and risk prediction error were scaled within the ambiguous condition. Both prediction errors were set to 0 in the control condition. Payoffs were scaled using data from the two conditions (regressors were scaled but not centered because the 0 is meaningful). The correlation between the payoff and the reward prediction error was large, $r = 0.49$, $p < 0.001$. The correlation between the reward and risk prediction error was negligible, $r = 0.00$, $p = 0.70$. Because none of the correlation was >0.75, multicolinearity was not an issue.

To illustrate what is going on, payoff, reward prediction error, and risk prediction error were plotted in two blocks of the ABCD and EFGH versions for subject 6. In the ABCD version (including big and infrequent losses), it appears that a big loss resulted in a negative reward prediction error and a positive risk prediction error (Fig. 5). In the EFGH version (including big and infrequent wins), it appears that a big win resulted in a positive reward prediction error and a positive risk prediction error (Fig. 6). In the control blocks, the reward and risk prediction errors are set to 0, but payoff values are still taken into account.

For the ROI analyses, the amplitude of brain activation was estimated for each trial with a fixed effect GLM in BrainVoyager. A

Predictors time line (ABCD)
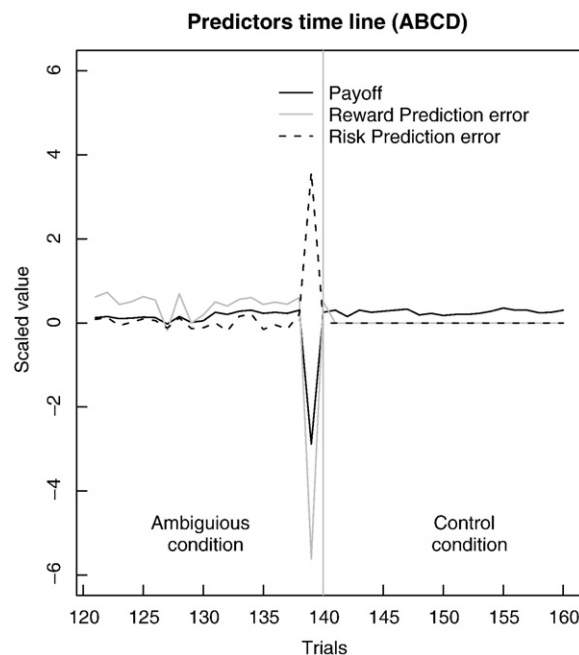
Fig. 5. Payoff, reward prediction error, and risk prediction error plotted as a function of trial in the ABCD version of the task for subject 6.

Fig. 6. Payoff, reward prediction error, and risk prediction error plotted as a function of trial in the EFGH version of the task for subject 6.

predictor was defined for each item, taking value 1 when the item occurred, and 0 otherwise. Predictors were convolved with the Hemodynamic Response Function (HRF). The beta values obtained from Brain Voyager were then analysed with the statistical program R, using the lme function (mixed linear model). The flexibility of R allowed us to explore the data extensively, to introduce several random factors, to analyse data separately for the ambiguous/control conditions, and to enter hemisphere (side) as an interaction effect. Two nested factors were entered in the random part of the mixed linear model in R: Block $\subset$ Subject. The first fixed regressor was the condition (0 for Control and 1 for Ambiguous). Exploratory analysis revealed that the first trial of each block provoked a greater brain activity. Therefore, it was entered as a second fixed regressor in the model (1 for first trial, 0 otherwise). Finally, reward prediction error ($\bar{\delta}$), risk prediction error ($\text{sign}(\xi) \cdot \sqrt{|\xi|}$), and payoff were entered in the model. Results revealed that the 2 random effects were significant for all regions of interest. This was also the case for the first trial of each block. Complete results of the mixed linear model for the insula are presented in Table 3. In ROI analyses, the α level was set at 0.001.

To test for laterality effects, interactions with the Side were entered in the model. None of the interaction effects appeared to be significant. Interactions with Gender were entered in separated analyses. Results revealed no significant interaction with gender.

Effects of the condition, risk prediction error, reward prediction error, and payoff are reported in Table 4 for all ROIs. Random and first

Table 3
Mixed linear model for the insula ROI.

| Variable	Estimate	Lower	Upper	Df	t	p
Random effect (SD)						
Subject	0.093*	0.035	0.244	–	–	–
Block	0.178*	0.154	0.205	–	–	–
Fixed effect						
Intercept	0.329*	0.210	0.447	–	–	
First trial	0.162*	0.120	0.204	12476	12.64	0.000
Payoff	−0.003	−0.015	0.009	12476	−0.84	0.401
Reward PE	0.006	−0.009	0.021	12476	1.39	0.164
Risk PE	0.024*	0.011	0.038	12476	5.84	0.000
Ambiguous condi.	0.004	−0.064	0.073	311	0.22	0.829

*0 not included in the 99.9% Confidence Interval ($p<0.001$).

trials effects are not shown. An important limit of the p-value is its dependence on the number of observations. One way to avoid this limitation is to assess the relative contributions of the predictors by comparing their associated t-values. To allow such a comparison, t-values were plotted for each ROI (Fig. 7). The region mostly related to the ambiguous conditions was the middle frontal gyrus ($t = 5.62$, $p<0.001$) followed by the lateral orbitofrontal cortex ($t = 3.44$, $p<0.001$). For reward prediction error, the strongest effect was found in the ventral striatum ($t = 7.46$, $p<0.001$), then in the dorsal striatum ($t = 5.72$, $p<0.001$). For risk prediction error, the strongest effect was found in the inferior frontal gyrus ($t = 6.04$, $p<0.001$) and

Table 4
Mixed linear model for all ROIs.

Variable	Estimate	Lower	Upper	t	p
Anterior cingulate					
Payoff	0.004	−0.011	0.018	0.84	0.401
Reward PE	0.012	−0.005	0.030	2.35	0.019
Risk PE	0.019*	0.003	0.035	4.00	0.000
Ambiguous condi.	−0.003	−0.081	0.075	−0.14	0.892
Posterior cingulate					
Payoff	0.029*	0.008	0.051	4.55	0.000
Reward PE	0.014	−0.013	0.040	1.71	0.087
Risk PE	−0.002	−0.026	0.022	−0.27	0.788
Ambiguous condi.	0.038	−0.077	0.153	1.10	0.270
Dorso medial PC					
Payoff	0.002	−0.012	0.015	0.42	0.678
Reward PE	0.018*	0.002	0.035	3.61	0.000
Risk PE	0.006	−0.009	0.021	1.28	0.201
Ambiguous condi.	0.061	−0.018	0.141	2.56	0.011
Middle frontal gyrus					
Payoff	0.004	−0.012	0.019	0.77	0.444
Reward PE	0.024*	0.006	0.043	4.28	0.000
Risk PE	0.008	−0.009	0.025	1.56	0.118
Ambiguous condi.	0.144*	0.059	0.229	5.62	0.000
Inferior frontal gyrus					
Payoff	−0.014	−0.028	0.000	−3.24	0.001
Reward PE	0.019*	0.002	0.036	3.72	0.000
Risk PE	0.029*	0.013	0.044	6.04	0.000
Ambiguous condi.	0.051	−0.032	0.134	2.03	0.043
Insula					
Payoff	−0.003	−0.015	0.009	−0.84	0.401
Reward PE	0.006	−0.009	0.021	1.39	0.164
Risk PE	0.024*	0.011	0.038	5.84	0.000
Ambiguous condi.	0.004	−0.064	0.073	0.22	0.829
Lateral OF					
Payoff	0.005	−0.012	0.022	1.06	0.291
Reward PE	0.018	−0.003	0.039	2.87	0.004
Risk PE	0.012	−0.007	0.031	2.14	0.032
Ambiguous condi.	0.090*	0.003	0.177	3.44	0.001
Medial OF					
Payoff	0.017	−0.001	0.036	3.04	0.002
Reward PE	0.006	−0.017	0.029	0.91	0.364
Risk PE	−0.006	−0.027	0.015	−0.96	0.337
Ambiguous condi.	0.021	−0.091	0.132	0.62	0.533
Dorsal striatum					
Payoff	−0.000	−0.014	0.014	−0.07	0.945
Reward PE	0.029*	0.012	0.046	5.72	0.000
Risk PE	0.010	−0.006	0.025	2.06	0.039
Ambiguous condi.	0.033	−0.041	0.107	1.50	0.135
Ventral striatum					
Payoff	0.010	−0.007	0.028	1.99	0.047
Reward PE	0.047*	0.026	0.068	7.46	0.000
Risk PE	0.023*	0.004	0.042	3.98	0.000
Ambiguous condi.	0.007	−0.071	0.085	0.30	0.765
Amygdala					
Payoff	0.016	−0.005	0.036	2.45	0.014
Reward PE	0.003	−0.022	0.029	0.40	0.689
Risk PE	0.021	−0.002	0.045	3.01	0.003
Ambiguous condi.	−0.018	−0.098	0.061	−0.77	0.442
Hippocampus					
Payoff	0.012	−0.003	0.026	2.69	0.007
Reward PE	−0.000	−0.017	0.017	−0.05	0.962
Risk PE	0.009	−0.006	0.025	1.96	0.050
Ambiguous condi.	−0.033	−0.108	0.042	−1.44	0.151

*0 not included in the 99.9% Confidence Interval ($p<0.001$).

Contribution to activity in the ROI
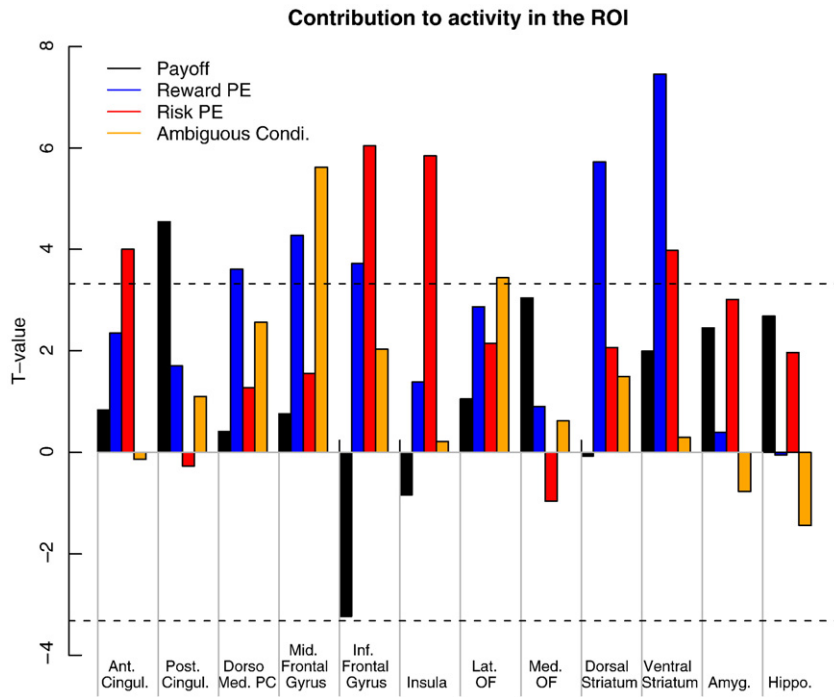
Fig. 7. Predictors *t*-value for each ROI. Dashed horizontal lines indicate the threshold at $\alpha = 0.001$, $df = 311$ (the df of the Condition was selected because it is the most conservative criterion).

the insula ($t = 5.84$, $p < 0.001$). The inferior frontal gyrus has the particularity to be negatively related to payoff and thus to encode losses ($t = -3.24$, $p = 0.001$). At $\alpha = 0.001$, it appeared that the dorsal striatum and the insula were uniquely related to reward and risk prediction errors respectively. The ventral striatum and the inferior frontal gyrus were related to both types of prediction errors.

Risk preference

In post-hoc analyses, we tested whether the two regions mostly related to risk prediction error were differently activated depending on participant's risk preference. To do so, brain activity in the inferior frontal gyrus and the insula was regressed on the risk preference parameter estimated from the reinforcement learning model. The α level was set to 0.01 because there were only 8 different observations (subjects) for risk preference. Results showed that risk aversion was associated with a higher activity in the inferior frontal gyrus in the ambiguous condition ($t = -4.07$, $p = 0.007$) but not in the control condition ($t = -2.36$, $p = 0.06$) (Table 5 and 6). This suggests that the inferior frontal gyrus is more active in risk aversive participants and preferentially in decision under ambiguity. The insula was neither related to risk preference in the ambiguous condition ($t = -1.95$, $p = 0.10$), nor in the control condition ($t = -1.24$, $p = 0.26$) (Table 7 and 8).

Discussion

The greatest effect found for making decision under ambiguity was observed in the middle frontal cortex, followed by the lateral orbitofrontal cortex. Impairment at the Iowa Gambling Task has been mostly observed in patients with lesion in the ventromedial prefrontal cortex, including varying sectors of the lateral orbitofrontal cortex (Bechara, 2004). This is compatible with the activation of the lateral orbitofrontal gyrus observed in the ambiguous condition. Interestingly, there exists also evidence of marked impairment at the Iowa Gambling Task following lesion of the dorsolateral prefrontal cortex (Manes et al., 2002; Clark et al., 2003), which fits with the activation of the middle frontal gyrus observed in the ambiguous condition. The orbitofrontal cortex is specialized in the rapid association between visual stimuli and reinforcers (Rolls et al., 1996). This associative mechanism is required in the Iowa Gambling Task because the deck values are unknown. As a consequence, it is necessary to associate the payoff with the deck selected just before. This is not necessary in the control condition since payoffs are written on decks, meaning that values are known in advance. The lateral prefrontal cortex has been implicated in the maintenance of information in working memory (Cohen et al., 1997). In the Iowa Gambling Task, the position of the advantageous decks needs to be maintained in working memory to allow the appropriate action. This is not useful in the control condition because the advantageous decks

Table 5
Inferior frontal gyrus ROI regressed on risk preference in the ambiguous condition.

Variable	Estimate	Lower	Upper	Df	t	p
Random effect (SD)						
Subject	0.059*	0.017	0.203	–	–	–
Block	0.209*	0.178	0.244	–	–	–
Fixed effect						
Intercept	0.466*	0.396	0.535	–	–	–
First trial	0.092*	0.042	0.142	6239	4.77	0.000
Risk preference	−0.110*	−0.210	−0.010	6	−4.07	0.007

*0 not included in the 99% Confidence Interval ($p < 0.001$).

Table 6
Inferior frontal gyrus ROI regressed on risk preference in the control condition.

Variable	Estimate	Lower	Upper	Df	t	p
Random effect (SD)						
Subject	0.103*	0.041	0.262	–	–	–
Block	0.220*	0.188	0.258	–	–	–
Fixed effect						
Intercept	0.396*	0.291	0.501	–	–	–
First trial	0.266*	0.209	0.322	6239	12.10	0.000
Risk preference	−0.096	−0.248	0.055	6	−2.36	0.056

*0 not included in the 99% Confidence Interval ($p < 0.001$).

position changes on every trial. The inferior frontal gyrus and the insula should also be critical for success in the Iowa Gambling Task as they appear to encode risk prediction error. Interestingly, patients with frontal lesions, including patients with lesions in the insula, have been found to make optimal investment decision in a situation where risk was associated with high expected return (Shiv et al., 2005). In this particular situation, impaired risk prediction constitutes an advantage.

Concerning payoff, results showed that the main effect of gains was located in the posterior cingulate cortex, followed by the medial orbitofrontal cortex. Losses were principally related to activity in the inferior frontal gyrus. McCoy et al. (2003) have shown by single neuron recording in monkeys that activity in the posterior cingulate gyrus was correlated to expected value associated with saccade and reward delivery. A common characteristic of the task used by McCoy et al. (2003) and the Iowa Gambling Task is that the location of the payoff is crucial in order to assess options utility. This may explain why both studies observed activity related to the posterior cingulate cortex. In human, reward delivery has been associated with activity in the medial orbitofrontal cortex (O'Doherty et al., 2003; O'Doherty et al., 2001; Knutson et al., 2001) as we did here with the Iowa Gambling Task. In addition, O'Doherty et al. (2003) found a positive correlation between loss and activity in the lateral orbitofrontal cortex. In the present study, this activation was also found in lateral regions, but in the inferior frontal gyrus. The representation of loss in the inferior frontal gyrus may serve to avoid bad choices because this region is implicated in inhibition (see below). One asset of our design and analyses, is that we were able to simultaneously take into account payoff and prediction errors. This was not done in previous studies, so that the payoff effect observed in the brain may in fact reflect reward or risk prediction errors.

Reward prediction error was mainly related to activity in the striatum, supporting numerous brain researches conducted in humans and monkeys. A study by Oya et al. (2005) has linked reward prediction error to neural activity in the paracingulate cortex during the Iowa Gambling Task. Neural activity was measured through deep implemented electrodes in a neurosurgical patient with behavioural performance in the normal range. However, due to the limited number of electrodes (3), the location of the peak activation associated with reward prediction error was unknown. Using fMRI, we extended these previous results by showing that the strongest effect of reward prediction error was located in the ventral striatum.

Based on results found under uncertainty and forced choice (Preuschoff et al., 2008), we formulated the hypothesis that risk prediction errors would be related to activity in the insula under ambiguity and free choice. To estimate risk under ambiguity, we used a risk-sensitive reinforcement learning algorithm. Risk prediction error derived from this algorithm was related to activity in insula, corroborating the hypothesis. In addition, a relationship between risk prediction error and activity in the inferior frontal gyrus emerged. These relationships held after controlling for payoff magnitude and reward prediction error. Interestingly, activities in the same three ROIs have already been reported in a reinforcement learning study (Li et al., 2006). The authors found a significant correlation between reward

Table 7
Insula ROI regressed on risk preference in the ambiguous condition.

Variable	Estimate	Lower	Upper	Df	t	p
Random effect (SD)						
Subject	0.061*	0.021	0.178	–	–	–
Block	0.174*	0.149	0.204	–	–	–
Fixed effect						
Intercept	0.338*	0.271	0.405	–	–	–
First trial	0.096*	0.053	0.140	6239	5.70	0.000
Risk preference	−0.050	−0.146	0.045	6	−1.95	0.099

*0 not included in the 99% Confidence Interval (p<0.001).

Table 8
Insula ROI regressed on risk preference in the control condition.

Variable	Estimate	Lower	Upper	Df	t	p
Random effect (SD)						
Subject	0.098*	0.041	0.238	–	–	–
Block	0.183*	0.156	0.215	–	–	–
Fixed effect						
Intercept	0.324*	0.226	0.421	–	–	–
First trial	0.228*	0.179	0.278	6239	11.86	0.000
Risk preference	−0.047	−0.187	0.093	6	−1.24	0.260

*0 not included in the 99% Confidence Interval (p<0.001).

prediction error and ventral striatum activity. They also found increased activities in the inferior frontal gyrus and the insula when reward structure changed unexpectedly, which increased variance in reward prediction error. The introduction of new reward structure was not captured by their reinforcement algorithm. Risk prediction error as modeled in the present study precisely tracks changes of the variance of reward prediction error. The use of a reinforcement model for risk allows us to put forward a new explanation: the inferior frontal gyrus and the insula encode changes in reward variance during reinforcement learning.

The functional difference between the inferior frontal gyrus and the insula remains to be understood. Previous studies have shown that the inferior frontal gyrus plays a crucial role in behavior inhibition (Aron et al., 2004). In the present study, the BOLD response of the inferior frontal gyrus was found to be more pronounced for risk aversive participants in the ambiguous situation. It is thus possible that the function of the inferior frontal gyrus is to inhibit risky choices. In line with our results, disruption of activity in the dorsolateral prefrontal cortex with repetitive transcranial magnetic stimulation induces risk taking behavior in risky (Knoch et al., 2006; Fecteau et al., 2007a) and ambiguous situation (Fecteau et al., 2007b). However, these results are at odds with the positive correlation found by Huettel et al. (2006) between posterior parietal activity and preference for risky over sure lotteries. It remains that the involvement of the inferior frontal gyrus in the implementation of an action policy to handle risk is supported by the fact that Preuschoff et al. (2008) found an activation related to risk prediction error in the insula but not in the inferior frontal gyrus. In the present study, participants were free to make choices whereas in Preuschoff et al. (2008) study, options selection was out of the participants' control. Future research is necessary to more directly test this hypothesis by manipulating experimentally the absence or presence of free choice and see if the inferior frontal gyrus encodes risk prediction error specifically when subjects make free choices.

If a risk alerting signal serves to inhibit risky choices in the inferior frontal gyrus, the signal in the insula may be more involved in the affective experience associated with risk. Anatomically, the insula conveys information from the body and it might serve as a gateway to the central nervous system for somatic reactions (Bechara and Damasio, 2005). Interestingly, skin conductance responses have been observed before the selection of risky decks in the Iowa Gambling Task, independent of the expected value of the decks (Tomb et al., 2002). These results combined with ours suggest that the insula may relay somatic response in reaction to increased risk. The involvement of the insula in risk learning is also compatible with the view that anxiety activates the insula. Indeed, authors have observed that insula activation after risky decision-making was more pronounced among participants with an anxious personality trait. Anxiety is also marked by a tendency to focus attention on bodily reactions and to overestimate the risk of failure (Butler and Mathews, 1983). Formulated within the somatic maker theory (Damasio, 1994), the function of the insula may be marking risky events with emotions so that they can benefit from a higher degree of relevance in future decisions. In line with this reasoning, it has been shown that

emotional events are better encoded and retrieved in memory (D'Argembeau et al., 2006) and these processes are supported by limbic structures including the insula (LaBar and Cabeza, 2006).

When modeling valuation under uncertainty, economists generally favor the expected utility theory. Finance academics and professionals, however, prefer to value risky prospects with the mean–variance model. Risk sensitivity is explained differently by the two approaches. Expected utility theory predicts that the preference for risk is related to the curvature of the utility function. The faster marginal utility decreases as a function of payoff, the higher the risk aversion. This strategy has been translated in reinforcement learning by applying a nonlinear transformation to individual rewards (Howard and Matheson, 1972) or to the reward prediction error (Mihatsch and Neuneier, 2002). As a result, the reinforcement learning algorithm becomes sensitive to risk, but without any measure of risk. In contrast, the mean–variance preference theory posits that risk aversion is the result of the penalty imposed on variance. This can be seen in our reinforcement learning algorithm where the utility of each deck is a linear combination of estimated expected value and variance. This implies that a measure of variance is computed separately from the expected value. To date it is unclear whether the human brain computes values in accordance with the expected utility theory or with the mean–variance analysis. If the first model is true, neuroscientists should only observe brain activities related to the expected utility of an option. If the second model is correct, they should see two signals in the brain: One for the expected value and the other for a risk measure like variance.

A central assumption of expected utility is also that the value of a prospect is computed by multiplying probabilities of each possible "state of nature" with the payoff utility in that state, and summing the results. Thus in ambiguous situations, state probabilities need to be learned, e.g. using Bayesian updating. For the mean–variance theory, uncertainty relies on payoff and learning focuses on moments (mean and variance). Based on this fundamental difference, recent behavioral results indicate that humans mix the two approaches when making decision under uncertainty, but rely more on mean–variance analysis when the number of state probabilities is high (d'Acremont and Bossaerts, 2008). In such a situation, estimating probabilities with Bayesian updating becomes unreliable (Diaconis and Freedman, 1986). Reinforcement learning is more suitable because state probabilities are not necessary to compute the expected value and variance. The Iowa Gambling Task represents a typical situation where probabilities cannot be tracked due to the high number of states. For instance, in Decks A, B, C, and D there are 31, 16, 22, and 18 different payoffs, respectively. So it is impossible to accurately estimate the probability of the occurrence of each payoff given the limited sampling at disposition (100 for 4 decks). In other words, if participants rely on expected utility in order to take risk into account, there is no reason to observe a neural signature of risk prediction errors (because variance does not need to be estimated). On the contrary, the use of reinforcement learning to estimate mean and variance predicts the neural signature of risk prediction error we observed.

In the Iowa Gambling Task, the selection of a stimulus is immediately followed by a reward as illustrated in Fig. 8 (left). In such a situation, the Rescorla–Wagner rule can be used to estimate the expected value and variance of the payoff that will be delivered in the next time step (Preushoff and Bossaerts, 2007) and this is the strategy followed in the present study. In a more natural setting, multiple stimuli generally overlap with multiple rewards as depicted in Fig. 8 (right). In this complex situation, Temporal Difference (TD) learning is able to estimate the expected value of the total reward (Sutton and Barto, 1998), but not its variance (total risk). It is only capable of predicting variance in the next time step (one-step ahead risk). Mathematical developments and simulations indicate that TD learning can be modified to evaluate the variance of the total reward
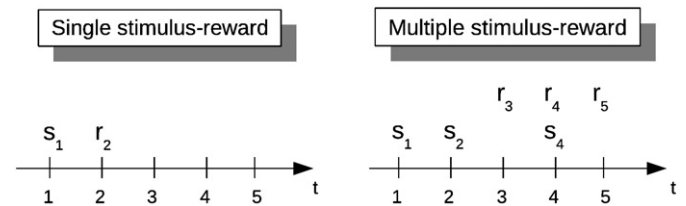

Fig. 8. Trial organization of single stimulus-reward (Left) and multiple-stimuli, multiple-rewards (Right) environment.

(d'Acremont et al., 2009). An important avenue for future research is to test whether one-step ahead and total risk are related to distinct neural signatures.

To conclude, results derived from a novel reinforcement algorithm revealed that reward and risk prediction errors are processed by distinct regions in the human brain. More specifically, changing risk correlates with activity in the insula and the inferior frontal gyrus. The latter region is also more activated in risk aversive individuals during decision-making under uncertainty. Reward prediction error is related to activity in sub-cortical regions and risk prediction error to activity in cortical regions. From an evolutionary perspective, this may indicate that the development of risk learning emerged later in the phylogeny compared to reward learning. Neural activity in response to risk choice has been observed in two male rhesus macaques (McCoy and Platt, 2005), but the presence of risk prediction error remains unexplored in animals. It would be interesting in the future to explore whether risk prediction error is also part of reinforcement learning across species.

Acknowledgments

References

Anderson, S., Bechara, A., Damasio, H., Tranel, D., Damasio, A., 1999. Impairment of social and moral behavior related to early damage in human prefrontal cortex. Nat. Neurosci. 2, 1032–1037.

Aron, A., Robbins, T., Poldrack, R., 2004. Inhibition and the right inferior frontal cortex. Trends Cogn. Sci. 8, 170–177.

Bechara, A., 2004. The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage. Brain Cogn. 55, 30–40.

Bechara, A., Damasio, A., 2005. The somatic marker hypothesis: a neural theory of economic decision. Games Econom. Behav. 52, 336–372.

Bechara, A., Damasio, A., Damasio, H., Anderson, S., 1994. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50, 7–15.

Bechara, A., Damasio, H., Tranel, D., Damasio, A., 1997. Deciding advantageously before knowing the advantageous strategy. Science 275, 1293–1295.

Bechara, A., Tranel, D., Damasio, H., 2000. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. Brain 123, 2189.

Bell, D.E., 1995. Risk, return, and utility. Manag. Sci. 41, 23–30.

Butler, G., Mathews, A., 1983. Cognitive processes in anxiety. Adv. Behav. Res. Ther. 5, 51–62.

Clark, L., Manes, F., Antoun, N., Sahakian, B., Robbins, T., 2003. The contributions of lesion laterality and lesion volume to decision-making impairment following frontal lobe damage. Neuropsychologia 41, 1474–1483.

Cohen, J., Perlstein, W., Braver, T., Nystrom, L., Noll, D., Jonides, J., 1997. Temporal dynamics of brain activation during a working memory task. Nature 386, 604–608.

d'Acremont, M., Bossaerts, P., 2008. Neurobiological studies of risk assessment: a comparison of expected utility and mean–variance approaches. Cogn. Affect. Behav. Neurosci. 8, 363–374.

d'Acremont, M., Gilli, M., Bossaerts, P. (2009). Predicting risk in a multiple stimulus-reward environment. In: Dreher, J.-C., Tremblay, L. (Eds.), Handbook of Reward and Decision Making. Academic Press.

Damasio, A., 1994. Descartes' Error: Emotion, Reason, and the Human Brain. Grosset/Putman, New-York.

Damasio, H., 2005. Human Brain Anatomy in Computerized Images. Oxford University Press, USA.

D'Argembeau, A., Van der Linden, M., d'Acremont, M., Mayers, I., 2006. Phenomenal characteristics of autobiographical memories for social and non-social events in social phobia. Memory 14, 637.

Diaconis, P., Freedman, D., 1986. On the consistency of Bayes estimates. Ann. Stat. 14, 1–26.

Engle, R., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.

Ernst, M., Bolla, K., Mouratidis, M., Contoreggi, C., Matochik, J., Kurian, V., 2002. Decision-making in a risk-taking task: a PET study. Neuropsychopharmacology 26, 682–691.

Fecteau, S., Knoch, D., Fregni, F., Sultani, N., Boggio, P., Pascual-Leone, A., 2007a. Diminishing risk-taking behavior by modulating activity in the prefrontal cortex: a direct current stimulation study. J. Neurosci. 27, 12500.

Fecteau, S., Pascual-Leone, A., Zald, D., Liguori, P., Theoret, H., Boggio, P., 2007b. Activation of prefrontal cortex by transcranial direct current stimulation reduces appetite for risk during ambiguous decision making. J. Neurosci. 27, 6212.

Fiorillo, C., Tobler, P., Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. Science 299, 1898–1902.

Friston, K., Holmes, A.P., Ashburner, J., 1999. Statistical Parametric Mapping (spm) [Computer software]. Available at http://www.fil.ion.ucl.ac.uk/spm/. Wellcome Department of Imaging Neuroscience, London, UK.

Fukui, H., Murai, T., Fukuyama, H., Hayashi, T., Hanakawa, T., 2005. Functional activity related to risk anticipation during performance of the iowa gambling task. NeuroImage 24, 253–259.

Galassi, M., Davies, J., Theiler, J., Gough, B., Jungman, G., Booth, M., 2006. GNU Scientific Library Reference Manual. Network Theory Limited, UK.

Howard, R., Matheson, J., 1972. Risk-sensitive Markov decision processes. Manag. Sci. 18, 356–369.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., Camerer, C., 2005. Neural systems responding to degrees of uncertainty in human decision-making. Science 310, 1680–1683.

Huettel, S., Song, A., McCarthy, G., 2005. Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. J. Neurosci. 25, 3304.

Huettel, S., Stowe, C., Gordon, E., Warner, B., Platt, M., 2006. Neural signatures of economic preferences for risk and ambiguity. Neuron 49, 765–775.

Knoch, D., Gianotti, L., Pascual-Leone, A., Treyer, V., Regard, M., Hohmann, M., 2006. Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. J. Neurosci. 26, 6469.

Knutson, B., Bossaerts, P., 2007. Neural antecedents of financial decisions. J. Neurosci. 27, 8174.

Knutson, B., Fong, G., Adams, C., Varner, J., Hommer, D., 2001. Dissociation of reward anticipation and outcome with event-related fMRI. NeuroReport 12, 3683.

Kuhnen, C., Knutson, B., 2005. The neural basis of financial risk taking. Neuron 47, 763–770.

LaBar, K., Cabeza, R., 2006. Cognitive neuroscience of emotional memory. Nat. Rev. Neurosci. 7, 54.

Li, J., McClure, S., King-Casas, B., Montague, P., 2006. Policy adjustment in a dynamic economic game. PLoS ONE 1, e103.

Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., 2002. Decision-making processes following damage to the prefrontal cortex. Brain 125, 624.

Markowitz, H., 1952. Portfolio selection. J. Finance 7, 77–91.

McClure, S., Berns, G., Montague, P., 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38, 339–346.

McCoy, A., Platt, M., 2005. Risk-sensitive neurons in macaque posterior cingulate cortex. Nat. Neurosci. 8 (9), 1220–1227.

McCoy, A., Crowley, J., Haghighian, G., Dean, H., Platt, M., 2003. Saccade reward signals in posterior cingulate cortex. Neuron 40, 1031–1040.

Mihatsch, O., Neuneier, R., 2002. Risk-sensitive reinforcement learning. Mach. Learn. 49, 267–290.

O'Doherty, J., Kringelbach, M., Rolls, E., Hornak, J., Andrews, C., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat. Neurosci. 4, 95–102.

O'Doherty, J., Critchley, H., Deichmann, R., Dolan, R., 2003. Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J. Neurosci. 23, 7931–7939.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R., 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304, 452–454.

Oya, H., Adolphs, R., Kawasaki, H., Bechara, A., Damasio, A., Howard, M., 2005. Electrophysiological correlates of reward prediction error recorded in the human prefrontal cortex. Proc. Natl. Acad. Sci. 102, 8351–8356.

Paulus, M., Rogalsky, C., Simmons, A., Feinstein, J., Stein, M., 2003. Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. NeuroImage 19, 1439–1448.

Preushoff, K., Bossaerts, P., 2007. Adding prediction risk to the theory of reward learning. Ann. N. Y. Acad. Sci. 1104, 135–146.

Preuschoff, K., Quartz, S., Bossaerts, P., 2008. Human insula activation reflects risk prediction errors as well as risk. J. Neurosci. 28, 2745–2752.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing (version 2.5.1) [Computer software]. Downloaded on http://www.r-project.org/. R Foundation for Statistical Computing, Vienna, Austria.

Rode, C., Cosmides, L., Hell, W., Tooby, J., 1999. When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory. Cognition 72, 269–304.

Rolls, E., Critchley, H., Mason, R., Wakeman, E., 1996. Orbitofrontal cortex neurons: role in olfactory and visual association learning. J. Neurophysiol. 75, 1970–1981.

Rolls, E., McCabe, C., Redoute, J., 2007. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. Cereb. Cortex.

Rothschild, M., Stiglitz, J.E., 1970. Increasing risk: I. A definition. J. Econ. Theory 2, 225–243.

Schultz, W., Dayan, P., Montague, P., 1997. A neural substrate of prediction and reward. Science 275, 1593–1599.

Sharpe, W., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. J. Finance 19, 425–442.

Shiv, B., Loewenstein, G., Bechara, A., Damasio, H., Damasio, A., 2005. Investment behavior and the negative side of emotion. Psychol. Sci. 16, 435–439.

Sutton, R., Barto, A., 1998. Reinforcement Learning: An Introduction. MIT Press.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System: An Approach to Cerebral Imaging. Thieme Medical Publishers, New York.

Tobler, P., Fiorillo, C., Schultz, W., 2005. Adaptive coding of reward value by dopamine neurons. Science 307, 1642–1645.

Tomb, I., Hauser, M., Deldin, P., Caramazza, A., 2002. Do somatic markers mediate decisions on the gambling task? Nat. Neurosci. 5, 1103–1104.

Tobler, P., O'Doherty, J., Dolan, R., Schultz, W., 2007. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. J. Neurophysiol. 97, 1621.

van Honk, J., Hermans, E., Putman, P., Montagne, B., Schutter, D., 2002. Defective somatic markers in sub-clinical psychopathy. NeuroReport 13, 1025.

Weber, E., Shafir, S., Blais, A., 2004. Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. Psychol. Rev. 111, 430–445.

Winker, P., Gilli, M., 2004. Applications of optimization heuristics to estimation and modelling problems. Comput. Stat. Data Anal. 47, 211–223.