Full Length Article

Decision-making under risk: A graph-based network analysis using functional MRI

Ludovico Minati a,b,⁎, Marina Grisoli c, Anil K. Seth d,e, Hugo D. Critchley b,e,f

Introduction

Many aspects of everyday life such as crossing a road, choosing to eat, spending or saving money, confronting wild animals in hunter-gatherer societies, require the ability to take decisions effectively. Typical decisional situations are characterized by multiple options, each carrying potential rewards, risks and associated outcome probabilities. Financial gambles provide a powerful tool to investigate systematically how motivational decisions are reached in the presence of risk. For example, a basic gambling game can be proposed where individual gambles differ in values for potential loss (kLOS), potential win (kWIN) and outcome probability (pWIN). If one has to choose among such “mixed” gambles, or decide whether or not to take an individual gamble, the locally optimal strategy would be to calculate the expected value \( EV = k_{\text{WIN}} \times p_{\text{WIN}} - k_{\text{LOSE}} \times (1 - p_{\text{WIN}}) \) for each option and systematically choose the ones with the highest \( EV \) (Machina, 1982; von Neumann and Morgenstern, 1944). However, in the real world biases such as risk-aversion (i.e. potential losses are overrepresented relative to gains) and loss-aversion (i.e. large potential losses are preferentially avoided irrespective of overall value) exert powerful influences. Prospect theory is a prominent account of realistic decision making, in which deviations from purely utilitarian behavior such as risk aversion are represented as non-linear value functions. Particularly, prospect theory postulates that choosing under risk involves an initial “editing” step, wherein the apparent economic parameters (e.g., \( k_{\text{WIN}}, k_{\text{LOSE}} \) and \( p_{\text{WIN}} \)) are integrated into abstract value representations (e.g., \( EV \)), followed by an evaluation step that compares the \( EV \) of each option to a reference point...
determined by the EVs of all options available in a given situation (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992; Wilkinson, 2008).

Published neuroeconomic studies of motivational decision-making indicate that both the evaluation of different choice options, and the process of choosing between them, involve an interplay between associative regions, mainly within prefrontal cortex, and phylogenetically-older structures including striatum, insula, cingulate cortex and amygdala (Alexander and Brown, 2011; Andersen and Cui, 2009; Clark, 2010; Kable and Glimcher, 2009; Levine, 2009; Platt and Huettel, 2008; Rangel and Hare, 2010; Rushworth and Behrens, 2008; Trepel et al., 2005; Wallis and Kennerley, 2010; Wang, 2008). To a certain extent, it is possible to link subcomponents of decision-making processes to specific brain regions. In current views, the dorsal and ventral lateral prefrontal cortices (DLPFC, VLPFC) and parietal lobe are implicated in integrative computation of EV, whereas abstract value representation involves primarily ventromedial and dorsomedial prefrontal cortices (VMPPC, DMPPC) and adjacent orbitofrontal cortex (OFC). Ascending dopamine pathways, projecting to the striatum and medial PFC, are associated with the representation of reward through prediction error signals that support motivational learning. In parallel, the anterior insular and cingulate cortices, which are linked to generation and representation of emotional bodily arousal, are involved in the representation of risk and uncertainty (Alexander and Brown, 2011; Andersen and Cui, 2009; Clark, 2010; Critchley, 2005; Kable and Glimcher, 2009; Levine, 2009; Platt and Huettel, 2008; Rangel and Hare, 2010; Rushworth and Behrens, 2008; Trepel et al., 2005; Wallis and Kennerley, 2010; Wang, 2008).

Most published functional neuroimaging studies in this area are reliant on univariate analytic techniques to establish direct relationships between specific economic variables and regional brain activity. The approach provides valuable insights, but it treats each region in isolation rather than as a node in a dynamical network. In contrast, a recent alternative perspective is that observable behavior arises as an emergent property of multiple areas interacting as a network (e.g., Sporns, 2009; Stephan, 2004).

Animal neuro-histological studies, and more recent human in-vivo neuroimaging investigations of structural and functional connectivity, suggest that topological features inherent in the organization of cortical networks underpin the brain’s capacity to combine efficiently functional specialization and integration (Buckner et al., 2009; Hagmann et al., 2008; Sporns, 2009; Sporns and Tononi, 2001; Sporns et al., 2000; Sporns et al., 2004). Specifically, the density of node connections (referred to as “degree” in the context of graph theory) is distributed very inhomogeneously across the cortex, with a small number of well-defined, discrete areas displaying a disproportionately large number of links (Hagmann et al., 2008; Sporns, 2009; Wang et al., 2010; Zemanová et al., 2006). Such key areas are hypothesized to act as “hubs”, with the dual purpose of non-linearly integrating information from multiple afferent connections and of relaying signals, thus optimizing the wiring and metabolic costs of long-range interactions (Bassett and Bullmore, 2006; Bullmore and Bassett, 2011; Guimerà et al., 2005; Zamora-López et al., 2010). Individual cortical hubs represent “convergence-divergence” zones, which exchange information bi-directionally with earlier sensory and association cortices to encode specific combinatorial arrangements of knowledge (e.g., representing an object or an abstract situation; Meyer and Damasio, 2009). The collective action of cortical hubs leads to the emergence of segregated large-scale networks that correspond to behavioral abilities in specific neuropsychological domains (Mesulam, 1998).

Univariate analysis alone is unable to provide a comprehensive neural account of risky decision-making, which involves complex interactions between multiple cognitive processes (e.g., working memory, strategic and probabilistic reasoning, and the abstract representation of relative magnitudes) integrated with affective and physiological response. Within a realistic neuroeconomic framework, relevant economic parameters are encoded not only as the intensity of activity within specific regions but also at network level, as modulations of dynamic connectivity between cortical hubs.

Among multivariate techniques, psychophysiological interaction (PPI) tests for the changes in the (physiological) influence of one “source” region on the activity of a second “target” region as a result of modulation by an external (psychological) factor, such as an economic parameter in a decisional task. If such an interaction is shown, the inference is that “target” region functionally integrates neural information from the “source” region with the task-determined psychological context (Friston et al., 1997). Another technique for examining communication across brain regions is dynamic causal modeling (DCM), whereby brain networks are treated as dynamic systems that are subject to external perturbation. DCM is used to gain quantitative estimates of the magnitude and direction of causal interaction between brain regions, by representing neuronal activity as a hidden variable within a biophysical model of observable hemodynamic responses. One can use Bayesian comparisons of pre-specified DCMs to achieve the best model of parameter effects on inter-regional effective connectivity (Friston et al., 2003; Stephan et al., 2007). To date there are only a few published studies applying connectivity analyses to financial decision-making: During reward processing activity in the ventral striatum is closely mirrored by responses across the insula, amygdala, hippocampus and orbitofrontal cortex (Camara et al., 2008). During a study of ‘charitable giving’, VMPPC activity integrated inputs from the anterior insula and temporal cortex to reflect the amount donated (Hare et al., 2010). Altered ‘value-dependent’ coupling between sub-genual cingulate and amygdala is observed when monetary rewards are linked to physical pain (Park et al., 2011). Each of these studies used a small number of seed regions of interest, that were defined either on the basis of results from univariate analyses or a priori hypotheses.

A complementary approach is to parcellate all brain areas potentially involved in task performance and perform PPI analyses across all of them to identify networks of connections modulated by relevant economic parameters. This approach has potential advantages in that it embeds minimal assumptions and yields whole-brain effective connectivity graphs that can be evaluated using established topological metrics, supporting inferences on the integrative role of each region and enabling the discovery of the cortical hubs involved in decision-making (Bassett and Bullmore, 2006; Camara et al., 2008; Camara et al., 2009; Salvador et al., 2005; Zamora-López et al., 2010). For example, this type of graph-based network analysis applied to functional resting-state data highlights abnormalities in large scale dynamics in early-stage dementia and schizophrenia (see Wang et al., 2010 for a review). Graph theory provides a wide range of topological metrics characterizing individual nodes, three parameters being particularly relevant in this context: i) “degree”, representing the number of connections for a given node, ii) “betweenness centrality”, representing the number of shortest paths passing through a given node and iii) “clustering coefficient”, representing the density of connections across the network “neighbors” of a given node. On graphs of effective connectivity, nodes exhibiting comparatively large values of degree and betweenness centrality and small clustering coefficients are likely to act as hubs integrating information from a representation distributed over multiple areas (Bassett and Bullmore, 2006; Bullmore and Bassett, 2011; Newman, 2010; Sporns, 2009; Wang et al., 2010).

Here, we sought to gain more comprehensive understanding of the neural bases to risky decision-making by considering in detail the processing of gambles from a network perspective. Healthy participants performed a functional MRI study during which they evaluated and decided whether to accept or reject “mixed” gambles in the absence of feedback. Effective connectivity graphs were constructed from PPI analyses with the aim of identifying key integrative structures
supporting the representation of the economic parameters characterizing the task. Additionally, we tested i) if the main inferences drawn through PPI analysis on effective connectivity directionality are concordant with Bayesian comparisons of DCMs, ii) if the observed PPI effects are, in fact, artifactual consequences of different response non-linearities across brain regions, iii) if approximating EV by its sign function leads to results similar to those obtained for continuous EV, iv) if the observed PPI effects are also apparent when choice-related variance is removed and v) if similar PPI effects are observed for accepted and rejected gambles separately. For succinctness, these additional analyses were only performed for EV, as the strongest driver of decisional behavior (Kahneman and Tversky, 1979; Wilkinson, 2008).

Methods

Participants

Nineteen healthy subjects (8 females, age 36±7 years, formal education 16±3 years) participated in the experiment, which was conducted at the Fondazione IRCCS Istituto Neurologico Carlo Besta (FINCB) and formally approved by the local institutional review board (project no. fMRI-DM) and by the research governance and ethics committee of the Brighton & Sussex Medical School (BSMS; PhD project no. 10/056/MIN). After being informed of the study purpose, all participants provided written informed consent. No money or material compensation was provided, and in accordance with local regulations participants were not playing with real currency.

The dataset used for this study was originally acquired in the context of a univariate correlational investigation published separately (Minati et al., 2012); due to data storage problems (loss), data from 3 of the initial 22 participants could not be used. All participants were naive to the experiment, right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971) and without a history of neurological or psychiatric disorders. None was taking medication affecting the central nervous system.

Gambling task

Participants viewed a back-projection screen covering approximately one third of the visual field through a mirror mounted on the head coil. Each gamble consisted of three parameters, presented visually as white numbers (Arial font) on a black background: potential loss (hereafter referred to as $k_{\text{LOSE}}$, top left of the screen), potential win ($k_{\text{WIN}}$, top right) and probability of winning ($p_{\text{WIN}}$, bottom central; Fig. 1a). From these parameters, expected value was derived according to $EV = k_{\text{WIN}} p_{\text{WIN}} - k_{\text{LOSE}} (1 - p_{\text{WIN}})$; the average amount at stake, referred to with $(k_{\text{WIN}} k_{\text{LOSE}})$, was also considered to analyze the fMRI data. The task was written in Matlab (Mathworks Inc., Natick MA, USA) using the Cogent 2000 toolkit (Wellcome Laboratory of Neurobiology, UCL, London, UK).

Participants were requested to evaluate the gambles in terms of ‘gut feeling’ without attempting any mathematical computation, giving equal consideration to the three parameters. They were asked to accept or reject each gamble rapidly, within a maximum of 6 s, and informed that no feedback would be provided but the computer would keep track of their responses in a virtual bank-account, starting from a balance of zero and updated with the amount won or lost each time. They were also made aware that the outcomes were not pre-determined, but would be decided by the computer each time the response button was pressed. They were encouraged to earn as much money as possible by deciding as though they were playing with real currency. Prior to scanning, all participants practiced on ten sample gambles.

Responses were made using an MRI-compatible keyboard according to a four-level scale, i.e. “confident reject” (left middle), “unsure reject” (left index), “unsure accept” (right index) and “confident accept” (right middle). Participants were asked to express their judgment using all four levels but told that there was no difference in the consequence of “confident” and “unsure” responses. This four-level scale was used to promote reflection on each gamble, rather than use of a fixed rule (Torn et al., 2007). Upon response, each gamble disappeared immediately and was replaced by a white fixation cross.

The gambles presented during each session were randomly extracted from a pool whose characteristics are shown in Fig. 1b–e. The range for $k_{\text{WIN}}$ and $k_{\text{LOSE}}$ was between 0€ and 300€, while $p_{\text{WIN}}$ was between 0.1 and 0.9. The distributions of $k_{\text{WIN}}$ and $k_{\text{LOSE}}$ were comparable (Mann–Whitney test $p = 0.5$) and the linear correlation coefficient with respect to EV was similar across $k_{\text{WIN}}$ ($r = 0.29$), $k_{\text{LOSE}}$ ($r = 0.31$) and $p_{\text{WIN}}$ ($r = 0.24$). This balancing ensured that the properties of the stimuli would not bias the evaluation toward one parameter over the others.

Each scanning session comprised four task blocks, each lasting 9 min. and followed by a pause of 3 min during which structural

Fig. 1. Stimuli and behavioral responses. a) Gamble presentation format. b–e) Distributions of $k_{\text{WIN}}$, $k_{\text{LOSE}}$, $p_{\text{WIN}}$ and EV in the gamble pool. f) Gamble acceptance probability (w) as a function of EV; the dotted lines denote 1 SD over participants.
imaging sequences were performed. Across the four blocks, 152 gamblings were presented with a constant inter-trial interval of 14.4 s. The first 4 gamblings after start and each pause were considered as warm-up trials and discarded.

**Data acquisition and preprocessing**

Imaging was performed at the neuroradiology department of the FINCB by means of a 1.5 T MR scanner (Magnetom Avanto, Siemens AG, Erlangen, DE), equipped with an 8-channel head matrix coil. Functional images sensitized to blood–oxygen level-dependant (BOLD) contrast were obtained using T2\textsuperscript{*}-weighted echo-planar sequences with TR = 2400 ms, TE = 45 ms, FoV 315 × 210 mm, matrix 90 × 60, 35 slices, 3.5 mm, no gap (isotropic voxel). For each task block, 240 volumes were acquired, yielding a total of 960 time points per scanning session. Corresponding phase/magnitude field-maps were acquired with a gradient-echo sequence having TR = 560 ms, TE = 4.7/9.5 ms. Anatomical scans were obtained using a volumetric T1-weighted magnetization-prepared rapid-acquisition gradient-echo sequence, with TR = 1640 ms, TE = 2 ms, FoV 256 × 256 mm, matrix 256 × 256, 160 slices, 1 mm, no gap.

The functional imaging data were pre-processed using SPM8 (Wellcome Trust Centre for Neuroimaging, UCL, London, UK) running under Matlab 7. All processing steps were completed on a Sun Blade 2500 workstation (Sun Microsystems Inc., Palo Alto CA, USA). After slice-timing correction, functional volumes were realigned within each session and un-warped using the corresponding field-maps. The average functional volume was co-registered with the individual anatomical scan utilizing an elastic transformation. The anatomical scan was thereafter segmented to obtain normalization matrices, which were subsequently applied to transform the functional volumes into MNI space. Spatial smoothing was then performed using an 8 mm FWHM kernel.

Subsequent processing steps were performed using software developed in-house. The time-courses of individual MNI-space voxels were filtered by removing the variance related to head translation and rotation using linear regression and by removing long-term fluctuations fitting the baseline with a third-degree polynomial and subtracting it.

**Behavioral response analysis**

The reaction times (RTs) were z-normalized at individual level, and trials for which they fell outside 3 SDs from the average or exceeded 6 s were removed. For each participant, the proportion of positive-EV gambles accepted and negative-EV gambles rejected, subsequently referred to as EV-discrimination, was calculated. The total amount earned, defined as the sum of the EVs of the accepted gambles, was also determined. To obtain a measure of risk aversion, two separate logistic regressions were performed modeling gamble acceptance as a function of $k_{\text{WIN}}$ and $k_{\text{LOSE}}$, and the regression coefficient ratio $\lambda = -\hat{p}_{\text{LOSE}}/\hat{p}_{\text{WIN}}$ was calculated (as in Tom et al., 2007). Further, individual decision-weighting curves representing gamble acceptance probability $w$ as a function of $EV$ were generated.

**Regions of interest (ROIs)**

On the basis of the literature on the functional neuroanatomy of financial decision-making (e.g. Kable and Glimcher, 2009; Platt and Huettel, 2008; Rangel and Hare, 2010; Trepel et al., 2005), the following regions were considered a priori to be of particular interest: primary motor cortex (limited to hand knob, MI), supplementary motor area (SMA), angular gyr (ANG), precuneus (PREG), early associative visual areas (OCC), anterior, middle and posterior (ACC, MCC and PCC) cingulate cortices, anterior insula (AINs), ventral-, dorsal- and antero-lateral prefrontal cortices (VLPFC, DLPFC and ALPFC), ventral- and dorsal-medial prefrontal cortices (VMPFC and DMPFC), orbitofrontal cortex (OFC), caudate nucleus (CAUD), lenticular nucleus (LENt), ventral striatum (VS), thalamus (THAL), amygdala (AMYG), hippocampus and parahippocampal gyrus (HPP).

These ROIs were initially defined on the basis of the automated anatomical labeling atlas (Tzourio-Mazoyer et al., 2002; Fig. 2) and manually refined to closely match anatomical boundaries on the high-resolution canonical brain provided in SPM8 by two experienced operators (LM and MG) using the MRlcr software (GSU/GT Center for Advanced Brain Imaging, Atlanta GA, USA). The Niewenhuys atlas was utilized as principal reference on gyral anatomy (Niewenhuys, 1996). Cingulate cortex was subdivided according to the areas defined by Vogt (2005). The prefrontal cortex was parcelled considering the criteria described in Rajkowska and Goldman-Rakic (1995), Fuster (1997) and Barbey and Graftman (2011).

**Response extraction and univariate analysis**

Regional time-courses were extracted by averaging over all voxels within each ROI having relative gray matter content exceeding 50%, as determined by the segmentation at pre-processing. The time-courses were epoched between 2.4 s (1 TR) before and 14.4 s (6 TRs) after each gamble onset, and the relative blood oxygenation level dependent signal percent change ($\Delta$BOLD%) derived with respect to the pre-stimulus baseline. For each gamble, the response amplitude was measured by averaging the $\Delta$BOLD% between 4.8 s and 9.6 s post-stimulus, corresponding to the peak of the hemodynamic response (Fig. 3). The response amplitudes were z-normalized across trials, for each region in each participant and the effect of variation in RTs was removed by linear regression.

As a first analysis, we sought to determine the direct effect of the economic parameters ($k_{\text{WIN}}, k_{\text{LOSE}}, p_{\text{WIN}}, EV$ and $\langle k_{\text{WIN}}, k_{\text{LOSE}} \rangle$) on regional responses and performed linear regressions over the pooled z-normalized response amplitudes of all participants, for each combination of parameter and region. We also determined whether, irrespective of the economic parameters, regional responses were significantly different from zero by means of one-sample $t$-tests. We accounted for multiple comparisons over the 42 regions controlling the family-wise error (FWE) by applying Bonferroni–Holm’s correction (Holm, 1979).

**Psychophysiological interaction (PPI) modeling**

We next applied PPI analysis to search for connections between pair of regions modulated by the economic parameters. In this analysis, a separate multilinear regression was performed for each parameter and pair of putative “source” and “target” regions, according to $y = \beta_0 + \beta_1 \cdot k + \beta_2 \cdot x + \beta_3 \cdot x \cdot k$, where $x$ and $y$ represent the z-transformed response amplitudes of the “source” and “target” regions respectively, $\beta_0$ is a constant term, $\beta_1$ corresponds to the direct effect of the economic parameter $k$, $\beta_2$ corresponds to the direct effect of activity in the “source” region (i.e., intrinsic functional connectivity) and $\beta_3$ corresponds to the interaction between the parameter $k$ and activity in the “source” region (i.e., effective connectivity).

Adjacency matrices for the connections modulated by each economic parameter ($k_{\text{WIN}}, k_{\text{LOSE}}, p_{\text{WIN}}, EV$ and $\langle k_{\text{WIN}}, k_{\text{LOSE}} \rangle$) were constructed by thresholding the significance of the $\beta_3$ term at $p<0.05$ FWE-corrected with Bonferroni–Holm’s correction over the 1722 (i.e., $42 \times (42 – 1)$) possible directed region pairs. To reduce connection clutter for graph visualization purposes, adjacency matrices were also generated for $p<0.01$ FWE.

Directed graphs of effective connectivity for each parameter were drawn from the adjacency matrices using the “Force Atlas” algorithm implemented in the Gephi 0.8 program (Bastian et al., 2009). For each node, the “degree” (i.e., total number of connections), “indegree” (i.e., number of “afferent” connections) and “outdegree” (i.e., number of
“efferent” connections) were computed (e.g., Bullmore and Bassett, 2011; Newman, 2010). The “betweenness centrality”, defined as the number of shortest paths passing through a node as given by Dijkstra’s algorithm, was also calculated, taking into account connection directionality (Brandes, 2001). Further, the “clustering coefficient” was obtained by triangle computations, taking into account connection directionality (Latapy, 2008).

EV was observed to be the strongest driver of behavior and the parameter for which the largest number of effective connections was observed. We therefore tested by means of a one-sample t-test whether the corresponding effective connectivity network for EV as a whole had small-world features, i.e., larger average clustering coefficient than 20 corresponding Erdös–Rényi random graphs, matched for number of connections, nodes and average path length. The small-worldness coefficient $S$ was computed to quantify the topological effect of clustering with respect to random connections (Humphries and Gurney, 2008). We also tested for significant modularity, i.e., presence of distinct network sub-clusters,

![Regions-of-interest (ROIs) used for signal extraction. The co-ordinates are given in MNI space. See Methods section for list of abbreviations.](image1)

![ΔBOLD% responses evoked by gamble presentation, averaged across all participants and stimuli. The error bars denote 1 SD over participants.](image2)
using algorithms described in detail elsewhere (Blondel et al., 2008).

To deal with the possible emergence of artefactual PPIs from different response non-linearities across regions, we repeated the PPI analysis removing its direct effect on responses in “source” and “target” regions up to the third degree. In other words, we considered $\hat{y} = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$, where $x = \alpha_1 EV - \alpha_2 EV^2 - \alpha_3 EV^3$, and similarly for $\hat{y}$. Further, given that gamble acceptance increases very rapidly in the $EV \approx 0$ region (see Results section), we hypothesized that the effect of $EV$ could be similar to that of a step transfer function. We therefore repeated the PPI analysis approximating $EV$ with its sign function and compared the resulting matrices of the $F$-values for the $\beta_3$ term across pairs of regions, representing these as heatmaps.

To investigate to what extent the observed PPI effects for $EV$ are related to value computation and representation rather than choice, we repeated the PPI analysis removing, for each “target”—“source” pair, the direct effect of choice on the “target” region and its interaction with the response of the “source” region. Further, we repeated the PPI analysis considering separately accepted and rejected gambles.

**Dynamic causal modeling (DCM)**

An additional analysis was performed to confirm the concordance of the directionality inferences drawn through PPI with respect to DCM10 (Wellcome Trust Centre for Neuroimaging, UCL, London, UK). Separate DCM analyses were undertaken for the five most densely interconnected regions on effective connectivity for $EV$ (bilateral ALPFC, bilateral DMPFC, right VMPFC; see Results section), modeling directed links with their neighbors within the prefrontal cortex. For each ROI, the first eigenvariate of the time-course was extracted and individually adjusted for the main effect of stimulus presentation as well as for the nuisance movement regressors. Individual bilinear, deterministic one-state DCMs (Stephan et al., 2007) were then specified for each target region. Two models were constructed and contrasted for each region: the first modeled afferent connections in the direction inferred from PPI analysis and the second modeled the connections in the opposite direction. For both models, $EV$ was entered as modulatory parameter for all connections and directly influencing the target region. For this part, gamble presentation was entered as ‘mini-blocks’ with duration corresponding to the measured RTs. After estimation at the level of individual participants, the models were entered in a random-effects Bayesian model selection analysis, which yielded group-level model exceedance probabilities for each pair of alternative models (Stephan et al., 2009).

**Results**

**Behavioral responses**

Participants discriminated positive and negative $EV$ beyond chance (78 ± 8% vs. 50%, one-sample $t$-test $p < 0.001$) and all earned a positive amount of ‘virtual’ money (1650 ± 540€ vs. 0€, one-sample $t$-test $p < 0.001$). The reaction times were 2.5 ± 0.9 s, and the 6 s limit was never exceeded. The regression coefficient ratio $\lambda$ indicated significant risk-aversion (1.3 ± 0.5 vs. 1, one-sample $t$-test $p = 0.02$). Gamble acceptance probability increased monotonically with $EV$ (Fig. 1f).

**Univariate analysis**

The task evoked positive responses bilaterally across the majority of regions, i.e. MI, SMA, PREC, OCC, MCC, VLPFC, ALPFC, DLPFC, OFC, DMPFC, AINS, ANG, CAUD, LENT, THAL and VS. In the ACC, positive responses were detected only on the left hemisphere, in the VMPFC significant negative responses were detected bilaterally. In the PCC, AMYG and HPP there was no significant response (Fig. 3).

As expected, strong motor response-related effects were expressed in MI bilaterally as correlations with $k_{WIN}$, $k_{LOSE}$ and $EV$ (Fig. 4). Expected value ($EV$) was positively correlated with response amplitude in right OCC ($F = 18.8$), left ACC ($F = 10.6$), VMPFC bilaterally ($F = 9$ for right, $F = 9.9$ for left) and right DMPFC ($F = 9.6$); no negative correlations were observed for this parameter. Potential loss ($k_{LOSE}$) was negatively correlated with response amplitude across left LENT ($F = -18.0$), VMPFC bilaterally ($F = -17.7$ for right, $F = -19.4$ for left), left VS ($F = -13.5$), right DLPFC ($F = -12.3$), ACC bilaterally ($F = -9.1$ for right, $F = -12.2$ for left), left MCC ($F = -10.8$), right ANG ($F = -10.0$), left ALPFC ($F = -9.8$), right OCC ($F = -9.7$), right PREC ($F = -9$) and right PCC ($F = -8.8$); no positive correlations were observed for this parameter. No significant correlations were observed for $k_{WIN}$ and $p_{WIN}$ beyond MI. Average amount ($\langle k_{WIN}, k_{LOSE} \rangle$) was negatively correlated with response amplitude in right ANG ($F = -10.1$) and right DLPFC ($F = -10.4$).

**Connections modulated by $EV$**

An effective connectivity network was identified for the representation of expected value ($EV$), in which the right VMPFC, right DMPFC,
right ALPFC, left ALPFC and left DMPFC appeared as prominent hubs (Fig. 5 and Fig. 6). At the FWE-corrected threshold of \( p < 0.05 \) (adopted for all network metrics calculations), the completeness was 9.7%, i.e. 168 directed connections out of 1722 were significantly modulated by EV; at \( p < 0.01 \) (adopted for display purposes only to reduce connection clutter), the completeness was 7.1%, corresponding to 122 connections. At both thresholds, 84% of modulated connections were negative, i.e. had \( \beta_3 < 0 \).

The regions with highest degree (Fig. 7), i.e. above the 90th percentile, were the right VMPFC (30 connections), right DMPFC (30), right ALPFC (29), left ALPFC (21) and left DMPFC (21). The spread observed for indegree was considerably larger than for outdegree (SD 6.6 vs. 3.2), i.e. all regions had between 5 and 10 “efferent” connections, whereas a small number had a substantially larger number of “afferent” connections, up to 25. Specifically, the regions with highest indegree, above the 90th percentile, were the right DMPFC (25 “afferent” connections), right VMPFC (22), right ALPFC (21), left DMPFC (17) and left ALPFC (14).

Betweenness centrality was highest, i.e. above the 90th percentile, for the right ALPFC (141 shortest paths passing through), right DMPFC (85), right VMPFC (78), left SMA (47) and left VMPFC (44). Excluding regions with zero clustering coefficients, the lowest clustering values, i.e. below the 10th percentile, were observed in the right ALPFC (0.13 proportion of connections among neighbors), right VMPFC (0.16), left OCC (0.17), right DMPFC (0.19) and left ALPFC (0.19).

Overall, the small-worldness coefficient \( S \) was 6.2, indicating that the graph strongly exhibited the topological features of a small world network. Accordingly, the average clustering coefficient was substantially higher than that of comparable Erdős–Rényi random networks (0.37 vs. 0.06 ± 0.03, \( p < 0.001 \)) matched for mean path length (1.93 vs. 1.98 ± 0.21, \( p = 0.4 \)). The modularity coefficient was 0.14, indicating that the network was not significantly subdivided in subclusters (with respect to the accepted cut-off of 0.4; see Blondel et al., 2008).

There were strong correlations between the traditional PPI analysis, the analysis performed removing non-linear EV effects (\( r = 0.7, p < 0.001 \)) and the analysis performed approximating EV with its sign function (\( r = 0.89, p < 0.001 \)) when comparing the \( F \)-values for the modulation term \( \beta_3 \). Accordingly, the distribution of \( F \)-values across regions appeared to follow a very similar pattern for the three analyses (Fig. 8).

Further, there were strong correlations between the traditional PPI analysis, the analysis performed removing the direct and interaction effects of choice (\( r = 0.68, p < 0.001 \)) and the analysis limited to rejected gambles (\( r = 0.45, p < 0.001 \), even though the corresponding interaction \( F \)-values were diffusely smaller. Contrariwise, there was no significant correlation with the interaction \( F \)-values computed over accepted gambles, the majority of which were in the vicinity of zero, except for degenerate values along the diagonal (Fig. 9).

DCM revealed that, for each of the five hub regions under consideration (Fig. 10), the exceedance probability was significantly higher for the model assuming afferent connections as discovered by the PPI analysis than for the model in which the direction of these connections was inverted, i.e. 0.88 vs. 0.12 for left ALPFC, 0.88...
Connections modulated by the other economic parameters

A smaller effective connectivity network was discovered for potential loss ($k_{\text{LOSE}}$, Fig. 11), in which the majority of connections overlapped corresponding the $EV$-sensitive network (31 out of 38, 82%). In this simpler network, the right ALPFC and right VMPFC remained as hubs. At the FWE-corrected threshold of $p<0.05$, the completeness was 22%, i.e. 38 directed connections out of 1722 were significantly modulated by $k_{\text{LOSE}}$; at $p<0.01$, the completeness was 1%, corresponding to 17 connections. At both thresholds, the majority of modulated connections were positive, i.e. had $\beta_3 > 0$ (86% at $p<0.05$, 94% at $p<0.01$). The regions with highest degree, i.e. above the 90th percentile, were the right ALPFC (14 connections), left ALPFC (10), right DLPFC (6) and right VMPFC (6). The regions with highest indegree were the right ALPFC (11 “afferent” connections), left ALPFC (8), right VMPFC (6) and left ACC (3). Since the path length and neighbor connectivity are not representative for such small networks (e.g., Newman, 2010), other network metrics were not computed.

No significant effective connectivity was observed for potential win ($k_{\text{WIN}}$), probability of winning ($p_{\text{WIN}}$) or average amount magnitude ($\langle k_{\text{WIN}}, k_{\text{LOSE}} \rangle$) was observed.

Discussion

The present experiment was motivated to extend our understanding of the processing of individual “mixed gambles”, i.e. choice options presenting a potential reward associated with a risk, beyond the representation of individual economic parameters in regional responses to the level of network dynamics. From a neurom economics perspective, detailed analysis of the processing of simple, individual monetary gambles is highly relevant, as it enables the isolation of primary neural correlates of adaptive action selection behaviors from complicating factors such as uncertainty (dealing with incomplete information) and ambiguity (ignorance of outcome probabilities), reward prediction error signaling and learning (such as found in the presence of outcome feedback), and bargaining and social behavioral (such as found when playing with an opponent). Further, extending neuroeconomic investigations beyond a purely localizationist approach is necessary from a theoretical viewpoint, considering that decision-making behaviors are highly abstract and multifactorial, and therefore unlikely to be adequately explicable in terms of activity taking place with isolated regions. Graph-based analysis of effective connectivity networks explicitly addresses two related issues; i) whether the neural representation of value and other economic parameters is widely distributed or limited to interactions among a restricted set of regions and; ii) whether the regions where stimulus value is reported to correlate strongly with response amplitude (e.g. ventromedial PFC and striatum) also stand out as information-integrating hubs during a decisional task.

Our mixed-gambling task successfully reproduced the key behavioral economic observation of risk aversion: participants overemphasized potential losses relative to equivalent gains. The effect was smaller in comparison with scenarios where there is substantial economic involvement (i.e., logistic regression coefficient ratio $\lambda = 1.3$ when compared to $\lambda = 2$ in Tversky and Kahneman, 1992; Abdellaoui et al., 2007; Tom et al., 2007). However the reproducibility of risk aversion was reassuring in that participants were not simply approximating a mathematical computation, but applying similar decisional criteria to those that lead to the adaptive biases observed in real-world decision-making (Kahneman and Tversky, 1979; Wilkinson, 2008). Indeed, previous behavioral economic studies...
demonstrate that significant deviations from locally optimal behavior can be obtained in the absence of real monetary endowment, and that the availability of small rewards may even elicit paradoxical effects (Gneezy and Rustichini, 2000).

Consistent with an earlier traditional voxel-wise appraisal of this dataset using univariate Gaussian-field inference methods (Minati et al., 2012), the univariate analysis of average ROI responses presented here revealed representation of expected value (EV) in the amplitude of regional responses within early associative visual areas and ventral and dorsal medial prefrontal cortices (Fig. 4). Our findings here add to the evidence for prefrontal “gain brain” circuits that track potential gains, and in which medial prefrontal cortex harbors abstracted representations of value and potential reward across classes of actions and stimuli (e.g., Alexander and Brown, 2011; Chib et al., 2009; Gläscher et al., 2009; Smith et al., 2010; Wunderlich et al., 2009). Significant correlations between economic parameters and visual activity are noteworthy, and plausibly reflect back-projected effects of salience from prefrontal, cingulate, amygdalar and brainstem regions (Critchley, 2005; Serences, 2008; Vuilleumier et al., 2004; Wendt et al., 2011).

\[ y = \beta_0 + \beta_1 \cdot EV + \beta_2 \cdot x + \beta_3 \cdot EV \cdot x \]

\[ \hat{y} = \beta_0 + \beta_1 \cdot EV + \beta_2 \cdot \hat{x} + \beta_3 \cdot EV \cdot \hat{x} \]

\[ y = \beta_0 + \beta_1 \cdot EV + \beta_2 \cdot |EV| + \beta_3 \cdot \frac{EV}{|EV|} \cdot x \]

Fig. 7. Network statistics for inter-regional connections modulated by EV, calculated at \( p_{\text{FWE}} < 0.05 \). Indegree: number of “afferent” connections; outdegree: number of “efferent” connections; degree: overall number of connections; betweenness centrality: number of shortest paths among all possible nodes that pass through a given node; clustering coefficient: density of connections among the neighbors of a given node.

Fig. 8. Comparison of F-values for the modulatory effect of EV under i) traditional PPI analysis, ii) PPI analysis in which the effects of EV on activity in the source and target regions have been removed up to the third degree (i.e., \( \hat{x} = x - \alpha_0 - \alpha_1 \cdot EV - \alpha_2 \cdot EV^2 - \alpha_3 \cdot EV^3 \)), iii) PPI analysis where continuous EV was replaced by its sign function. Blue denotes negative, red positive sign (range ±50). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Our network-discovery analysis for effective connectivity revealed a network of striking richness (Fig. 5). While only five regions displayed correlation between EV and response amplitude, EV modulated effective connectivity between 36 regions (when stringently correcting for multiple comparisons over the 1722 directed connections rather than the 42 nodes). One immediate inference is that the representation of value is deeply embedded within dynamical interactions among brain regions, rather than only in the intensity of local activity. Since effective connectivity is conceptually distinct and mathematically independent from direct parametric correlation (Friston et al., 1997; Friston et al., 2003; Stephan et al., 2007), it is interesting to identify regions expressing their contributions to the encoding of EV differentially, i.e. where there is uncoupling of regional activity and network contribution. Anterior–lateral prefrontal cortex exemplifies this: direct correlations between response amplitude and EV were clearly absent ($p = 0.7$ for right, $p = 0.4$ for left) yet the number of significant EV-modulated connections was very large (29 for right, 21 for left, both above the 90th percentile of the degree distribution).

In recent years, it has been consistently observed that self-organized biological, social and technological networks display small-world topology. This suggests that small-worldness is a pervasive ecological plausible feature of dynamical collective behaviors (Humphries and Gurney, 2008). Our finding of small-world topology within brain systems supporting adaptive decision-making extends such observations to the level of the neurobiological mechanisms underlying the motivational, microeconomic behavior of the individual. Small-worldness generally represents an optimum trade-off between local and global connectivities, providing efficient information transfer with limited connectivity complexity (Bassett and Bullmore, 2006; Bullmore and Bassett, 2011; Ginestet et al., 2011; Newman, 2010; Watts and Strogatz, 1998). More specifically, simulations of neuronal systems demonstrate that small-world networks display dynamical properties that are very different from corresponding random assemblies, notably drastically increased long-range synchronizability of activity and the emergence of specific attractors (Barahona and Pecora, 2002; Bullmore and Bassett, 2011; Ginestet et al., 2011; Newman, 2010; Watts and Strogatz, 1998). We speculate that, in this context, small-worldness emerged from the evolutionary optimization of information transfer across regions involved in adaptive survival-related behavior. This view is strengthened by the relative lack of modularity, suggesting that the distributed representation of value is not built around specific sub-modules implementing segregable operations.

For each node, we computed multiple metrics representing independent features representing its local connectivity (degree, indegree

**Fig. 9.** Comparison of F-values for the modulatory effect of EV under i) traditional PPI analysis, ii) PPI analysis in which the direct and interaction effects of choice (accept or reject) on activity in the target region have been removed, iii) traditional PPI analyses performed separately for accepted and rejected gambles. Blue denotes negative, red positive sign (range ±50). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
and outdegree), its position in the directed network (betweenness centrality) and the structure of connections across its neighbors (clustering coefficient). These node parameters implicate the dorso-medial prefrontal cortex, right ventro-medial prefrontal cortex and the anterior–lateral prefrontal cortex as key network hubs (Figs. 6 and 7). In the context of structural and functional brain connectivity, large degree, large betweenness centrality and small clustering coefficient are the hallmarks of integrative cortical nodes (Buckner et al., 2009; Hagmann et al., 2008; Sporns, 2009; Sporns and Tononi, 2001; Sporns et al., 2000, 2004). In the context of international air traffic, nodes displaying these features are referred to as "connector hubs" and are highly critical to sustain efficient long-haul transfer of goods and passengers (Guimerà et al., 2005). Our findings provide novel evidence that advances our understanding of the neural substrates of risky decision-making by providing the first explicit demonstration of the integrative role of these regions at the level of effective connectivity. A substantial body of literature reporting parametric correlation between monetary and subjective value and responses in these areas already existed (e.g., Alexander and Brown, 2011; Andersen and Cui, 2009; Clark, 2010; Levine, 2009; Platt and Huettel, 2008; Rangel and Hare, 2010; Rushworth and Behrens, 2008; Trepel et al., 2005; Wallis and Kennerley, 2010; Wang, 2008), but remained open to the criticism that the observation of a local correlation could reflect a distal or a distributed process, rather than a primary causal relationship between regional activity and value representation processes (Ramsey et al., 2010). In other words, it could not be completely ruled out that correlations with medial prefrontal cortex activity could at least partly arise as a secondary expression of processing taking place elsewhere (Philiastides et al., 2010). Our results demonstrate that the medial and anterior lateral prefrontal cortices have a key integrative role, displaying converging effective connectivity involving nodes that do not even display significant activity correlations upon univariate analysis.

Hubs in dynamical networks may serve diverse functions, in spite of having similar topological properties of their connections. Notably, one can distinguish between hubs that mainly provide communicational facilities while performing limited local computation, and those that also contain significant processing and memory infrastructure (Barrat et al., 2008; Bullmore and Bassett, 2011; Newman, 2010; Sporns, 2009). Here, we are unable to definitively distinguish between these roles since no information is available about the temporal dynamics of inter-regional interactions, however we can hypothesize that processing nodes could also exhibit a direct...
correlation with EV due to their direct computational engagement, whereas purely communication nodes would be characterized by less functionally-specific activity and therefore not exhibit direct activity correlations. From this perspective, at the chosen significance thresholds the bilateral ventro-medial and right dorso-medial prefrontal cortices would be classified as processing nodes, whereas the left dorso-medial prefrontal cortex and the anterior–lateral pre-frontal cortex bilaterally would be classified as communication nodes. While this distinction must remain speculative for our data, particularly as it depends on the choice of significance thresholds, future research should explore whether the relationship between topological parameters and univariate task effects can identify different types of network hubs in fMRI analysis. As demonstrated by ex-vivo tracer studies, the ventral- and dorsal-medial prefrontal cortices have dense bi-directional connections with the orbitofrontal cortex, dorsal- and ventral lateral prefrontal cortex and anterior–lateral prefrontal cortex, and are also directly interconnected with the cingulate, hippocampus, amygdala and striatum (Barbas, 2000; Brog et al., 1993; Ongür and Price, 2000; Ongür et al., 2003). Anterior–lateral prefrontal cortex is densely interconnected with both dorsal–lateral and ventrolateral prefrontal cortices, as well as with subcortical structures (Barbeau et al., 2009). Recent meta-analyses of functional imaging studies propose that this region is organized in a hierarchical form, with the more anterior portion supporting more abstract integrative functions (Badre, 2005; Botvinick, 2005). Since the dorsal–lateral prefrontal cortex and ventral–lateral prefrontal cortex (e.g., Sakagami and Pan, 2007; Trepe et al., 2005) are involved in integrating multisensory information into abstract variables that subserve value computation, the hypothesis of a key integrative role of the medial and anterior prefrontal cortices accords with the organization of structural connections.

It is noteworthy that the dorsal and ventral striatum were in a marginal topological location in the effective connectivity network for expected value, i.e. were associated with low degree, zero indegree and zero betweenness centrality. This finding suggests that the striatum does not play a major integrative role in the evaluative aspects of risky decision-making embodied within this task. It is noteworthy also that a recent behavioral study performed with the same task stimuli demonstrated normal performance in patients with subcortical neurodegeneration (Minati et al., 2011). Our results therefore robustly associate primary value discrimination during choice-option evaluation with medial prefrontal cortical function. However, this notion is at odds with a recent study where correlations with EV and individual risk aversion were observed in the striatum (Tom et al., 2007). We speculate that this discrepancy could be due to the task design: in our study participants had to weigh potential losses and gains with their respective outcome probabilities, whereas in the study by Tom et al. (2007) the outcome probability was fixed at 50% and the decisional problem could therefore be reduced to a simple comparison between the two magnitudes. The need for more abstract processing here may have led to a shift toward predominantly cortical representation, reducing the engagement of basal ganglia circuitry. Also, it cannot be ruled out that the correlations observed by Tom et al. (2007) could have been secondary to reward expectation, since participants were remunerated with a fraction of their earning, whereas here no material compensation was available.

Lesion studies have convergently demonstrated that damage to the ventral (Bechara et al., 1999; Clark et al., 2008) as well as dorsal (Manes et al., 2002) mesial prefrontal cortex reduces the ability to discriminate between risk-advantageous and disadvantageous situations. Discrete damage to the orbitofrontal (Manes et al., 2002) and lateral prefrontal cortices (Clark et al., 2003; Manes et al., 2002) has less obvious effect on decision-making in the absence of uncertainty, i.e. when outcome probabilities are explicitly represented. Our results are in accord with this literature, indicating that the medial prefrontal cortex harbors key integrative hubs for the representation of expected value. We provide the additional insight that the anterior–lateral prefrontal cortex may also implement critical integrative functions, as indicated by the high density of connections. Further lesion studies are therefore motivated to disentangle the effect of focal caudal and rostral damage to the lateral prefrontal cortex, and to test the hypothesis that for a given volume rostral lesions will have a substantially larger impact on value representation. It is also important to consider that significant effects have also been observed for damage to the anterior insula and amygdala, which appeared rather peripheral in our reconstruction of the effective connectivity network (De Martino et al., 2010; Weller et al., 2009).

Importantly, throughout this study all PPI analyses were conducted on the basis of the directly-measured overall amplitude of individual hemodynamic responses, rather than using regressors convolved with the canonical function to disentangle overlapping activations; while significantly increasing task duration, this approach has the substantial advantage of removing assumptions of time-course consistency and linear summation, which have the potential to generate artifactual interactions in fast-paced designs where the responses to individual gambles overlap considerably (Duann et al., 2002; Serences, 2004; Wager et al., 2005). Further, our inferences on effective connectivity were drawn applying family wise error through an implementation of Bonferroni’s correction that very robustly controls for the probability of type I error (Holm, 1979). While we cannot exclude outright that our choice of ROIs could have introduced correlations between neighboring regions caused by overlapping activity, our network-level inferences are robust with respect to this potential source of error since the majority of connections were observed between clearly anatomically-separate areas. Since some of the regions for which we observed the most dense effective connectivity did not even display any significant parametric correlation between the economic parameters and regional response, generating the ROI set on the basis of a preliminary voxel-based univariate analysis would likely have concealed important network elements.

One cannot exclude a priori that non-linear correlations between EV (or generally any continuous task parameter) and response amplitude may be present, and non-linear effects are indeed well-documented in neuroeconomics (e.g., Hsu et al., 2009). This represents a potential problem for any kind of PPI analysis, because if a ‘target’ and a ‘source’ region exhibit activity correlating with EV with different degrees of non-linearity, the bilinear (β2) term of the PPI equation may become artifactually significant without a real underlying functional interaction. We addressed this problem by performing an additional analysis in which the direct effects of EV were removed up to the third degree separately from both ‘source’ and ‘target’ regions before modeling the functional interaction. The results endorsed those of the initial analysis, strengthening confidence in the observed patterns of effective connectivity (Fig. 8).

Notably, this is the first PPI study where the possibility of spurious interactions due to non-linear effects is explicitly addressed in this way.

We also gained further insight with respect to the coding of value by showing that the underlying effective connectivity pattern was mostly unaltered when the continuous EV was replaced with a sign function (positive or negative). This indicates that the network is principally engaged in the discrimination of risk advantageousness or disadvantageousness, reducing the continuum of abstract value to a more basic heuristic, i.e. comparing the weighted potential gain and loss (i.e., kWIN×p and kLOSE×(1−pWIN)) rather than summing them continuously (Fig. 8). Accordingly, we did not find any parametric modulation in response amplitude or connectivity for average amount magnitude (kWIN×kLOSE). A dichotomous comparison mechanism is also apparent within the behavioral data, where the acceptance probability function w(EV) rises steeply in the vicinity of EV=0 and is essentially flat elsewhere (Fig. 1f). Our results
therefore strongly endorse the view that the mesial prefrontal cortex acts as fundamentally a decisional comparator (Venkatraman et al., 2009; Wunderlich et al., 2009).

Inferences on directional connectivity can only be drawn indirectly from fMRI data, i.e. by considering the statistical relationships among the time series. Specifically, the directionality of effective connectivity in PPI analyses is extractable since the bilinear term for a connection between two regions is different from the same connection modeled in the opposite direction (Friston et al., 1997, 2003). DCM shares some conceptual similarities with PPI analysis, but is not equivalent as it embeds an explicit biophysical model of neurovascular coupling and relies on a Bayesian approach for comparing competing models (Stephan, 2004). We applied DCM to a subset of frontal regions for the purpose of confirming our inferences based on PPI analysis. Modeling afferent connections into the medial and anterior–lateral prefrontal cortices carried much stronger evidence compared to models assuming the opposite directionality, reassuring on the validity of our PPI observations (Fig. 10). An alternative method for confirming PPI network discovery results would be to use Granger causality, a data-driven method based on the relative predictability of time-series in the context of linear autoregressive modeling. While this framework is powerful and makes fewer assumptions on underlying causal structure as compared to DCM, it is best suited to steady-state data generated in the absence of modulating variables (such as EV) rather than event-related data incorporating modulating factors; furthermore, its application to fMRI is still controversial due to several factors including the potentially confounding effect of hemodynamic responses and the slow sampling rate (TR) (Bressler and Seth, 2011; Valdes-Sosa et al., 2011). Although for these reasons we did not apply Granger causality in the present study, in future work we will examine whether Granger causality can reveal additional features of causal connectivity in our data.

Univariate correlations and effective connectivity analyses indicated that potential gain (kLOSE) and outcome probability (pVM) are not significantly represented in BOLD activity as individual parameters (Fig. 4). However, since both are necessary for the calculation of EV, they must be reflected in neural activity at some level; a plausible interpretation of our finding is that their coding is subter in comparison to potential loss (kLOSE) and does not generate sufficient fluctuations in the intensity of post-synaptic activity to reflect on the detectable BOLD signal. By contrast, significant negative correlations between kLOSE and response amplitude were observed across the medial and lateral prefrontal cortices, cingular cortex, occipital–parietal cortex and basal ganglia. This effect goes along the “gain brain” notion that potential losses are encoded as diffuse deactivations but, remarkably, in comparison with the previous study by Tom et al. (2007) here the mismatch in the representation of potential losses and gains was substantially larger, with no positive correlation detected for potential gain in any region. We hypothesize that this result, which is in keeping with the previous voxel-based univariate analysis of this dataset (Minati et al., 2012), is due to the fact that our task demanded abstract integration of three economic parameters, rather than simple comparison of explicit potential loss and gain with equal probability. We further observed that potential loss modulated connectivity across a network largely representing a subset of the effective connectivity for EV, centered around the bilateral anterior–lateral prefrontal cortex and the right ventromedial prefrontal cortex (Fig. 11). This finding substantiates the view that the brain invests far more resources in the encoding of loss than potential gains, likely because in natural competitive games potential losses may prove fatal even when presented in the context of overall risk-advantageous choice options (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992; Wilkinson, 2008). An important issue with the design of the present study is that EV was strongly coupled with choice, and the observed connectivity modulations may therefore reflect deliberative behavior rather than primarily value encoding. We addressed this issue by performing an additional analysis in which choice-related effects and interactions were removed prior to inferring EV-sensitive connections. While the corresponding F-values were diffusely reduced, likely also as a consequence of removing a substantial part of EV-related variance that also correlated with choice, the overall connectivity pattern was significantly preserved (Fig. 9). To further explore the relationship between choice and EV representation, we also re-analyzed the data separating the accepted and rejected gambles, revealing that, albeit weaker, the connectivity pattern inferred with the overall PPI analysis is present over rejected gambles but clearly absent for accepted gambles. On one hand, the presence of EV-modulated connections in rejected gambles confirms that our initial, overall PPI inferences were not purely incidental to choice-related neural differences between accepted and rejected gambles. On the other, the absence of such effects in accepted gambles highlights a fundamental difference in value processing, which could speculatively be interpreted as evidence of deeper value encoding under risk-disadvantageous conditions. Due to the coupling between EV and choice, this effect cannot be fully resolved on the basis of our data alone. Further experiments should be conducted requesting participants to determine risk-advantageousness mentally without making any decision, and loosening the coupling between EV and choice for example by means of introducing trials where participants are requested to decide in a highly risk-averse or risk-prone manner (i.e., generally boosting or depleting the proportion of accepted gambles irrespective of EV).

Conclusion

This study provides the first comprehensive evaluation of effective connectivity during risky decision-making based on a network-discovery approach. Our findings reveal that expected value modulates effective connectivity across a network that includes and extends substantially beyond those nodes that express direct relationships between value and activity. This network is characterized by small-world topology, unitary architecture and prominently features medial and anterior–lateral prefrontal areas as key hubs, characterized by dense connectivity and high betweenness centrality. The differential expression of univariate task effects across these hubs suggests that the ventro-medial prefrontal cortex may particularly represent a processing hub, whereas the anterior–lateral prefrontal cortex may subserve primarily communication functions. Our findings thus extend existing literature by robustly demonstrating that the representation of value is dependent on distributed processing taking place across a widespread network, feeding information into a restricted set of integrative prefrontal regions. Additionally, dichotomizing the gambles into risk advantageous or disadvantageous revealed that the operation underlying the observed connectivity is more pertinent to comparison of weighted gains and losses than to summative determination of continuous value. Further analyses revealed that, albeit smaller, an effect of continuous value was present across rejected gambles, but clearly absent for accepted gambles. Our findings also demonstrate that potential loss is significantly over-represented in comparison with potential gain and outcome probability at the level of regional activity as well as in terms of effective connectivity. Since mapping effective connectivity revealed strongly significant effects involving regions for which univariate effects were clearly absent, the results of the present study have broad paradigmatic implications for neuroeconomics in general, demonstrating the value of multivariate analysis in obtaining a comprehensive account of the neural substrates of decisional processes.

Acknowledgments

LM was wholly funded and employed by the Fondazione IRCCS Istituto Neurologico Carlo Besta (FINCB) during the core period of
this study. This research also forms part of a BMS doctoral dissertation (10/056/MIN) and is based on preparatory work conducted while LM was previously employed by BSMS. Data acquisition was conducted at and supported by the FINCB. The authors are grateful to Francesca Epifani for excellent operational assistance during data acquisition and to Sylvie Piacentini for useful discussions. The results of this study have in part been presented at the “Mapping Functional Networks for Brain Surgery” ISMRM workshop (Milan, IT, Sep 2011).

AKS and HC gratefully acknowledge the support of the Dr. Mortimer and Theresa Sackler Foundation. AKS is also supported by the EPSRC fellowship EP/G007543/1. The authors are grateful to an anonymous reviewer for insightful feedback on the interpretation of network analysis results.

References


