Research report

Value and prediction error estimation account for volatility effects in ACC: A model-based fMRI study

Massimo Silvetti\textsuperscript{a,b,\ast}, Ruth Seurinck\textsuperscript{a,b} and Tom Verguts\textsuperscript{a,b}

\textsuperscript{a}Department of Experimental Psychology, Ghent University, Ghent, Belgium
\textsuperscript{b}GIfMI (Ghent Institute for Functional and Metabolic Imaging), Ghent University Hospital, Ghent, Belgium

\textbf{A B S T R A C T}

In order to choose the best action for maximizing fitness, mammals can estimate the reward expectations (value) linked to available actions based on past environmental outcomes. Value updates are performed by comparing the current value with the actual environmental outcomes (prediction error). The anterior cingulate cortex (ACC) has been shown to be critically involved in the computation of value and its variability across time (volatility). Previously, we proposed a new neural model of the ACC based on single-unit ACC neurophysiology, the Reward Value and Prediction Model (RVPM). Here, using the RVPM in computer simulations and in a model-based fMRI study, we found that highly uncertain but non-volatile environments activate ACC more than volatile environments, demonstrating that value estimation by means of prediction error computation can account for the effect of volatility in ACC. These findings suggest that ACC response to volatility can be parsimoniously explained by basic ACC reward processing.

© 2012 Elsevier Ltd. All rights reserved.

\textbf{1. Introduction}

In order to obtain resources for its preservation and achieve its behavioral goals, any organism has to establish an adaptive interaction with its environment. For example, an animal looking for food must discover which places in its environment are more likely to contain alimentary rewards. This class of problems is studied by reinforcement learning (RL), a theoretical framework developed in the field of machine learning (Sutton and Barto, 1998). The simplest way to maximize reward consists in estimating the probability with which a reward will follow some action or environmental state, and then selecting an action based on these estimates. For example, a weasel could have learned that a particular odor nearby a hole in the ground indicates the presence of a prey, and thus it could select the action of entering the hole. These estimates (or predictions; denoted by $V$ here) can be updated by evaluating how well they correspond to actually obtained rewards. The difference between predicted and actual rewards is called “prediction error” (here $d$). In other words, the estimates of $V$ are constantly updated by the estimated $d$.

Neurophysiological investigations showed that the anterior cingulate cortex (ACC) has neuronal populations that code for the predicted reward (Amiez et al., 2006; Matsumoto et al., 2003) and two different neural populations coding for prediction errors ($d$) (Amiez et al., 2006; Matsumoto et al., 2007).
of the $\delta$ populations codes for positive prediction errors (here called $\delta^+$), i.e., when the reward is bigger than expected, and the other for negative prediction errors (here called $\delta^-$), when the reward is smaller than expected. Prediction error-related activity in ACC has been confirmed in humans by EEG and fMRI studies (Jessup et al., 2010; Oliveira et al., 2007; Nunez Castellar et al., 2010).

Since many years, the ACC has been indicated as one of the main areas involved in adaptive control of behavior (Ridderinkhof et al., 2004), and operations other than reward prediction and prediction error have also been ascribed to it. For example, it has been proposed that ACC is involved in error processing (Critchley et al., 2005), in the estimation of the probability of committing an error (Brown and Braver, 2005), or in conflict monitoring (Botvinick et al., 2001). In earlier work we proposed a new neural model of ACC, the Reward and Prediction Model (RVPM) (Silvetti et al., 2011). This model is based exclusively on the available single unit evidence we described above. It computes reward expectations by means of prediction error signals. The RVPM has been able to provide a neurocomputational account of various EEG and fMRI findings in humans, including error detection, error likelihood estimation, and response conflict (Silvetti et al., 2011), showing how heterogeneous experimental findings on the ACC could be interpreted as deriving from the computation of reward expectations and prediction errors.

Besides the proposed ACC functions cited above, a recent and influential paper showed that volatility modulates ACC activity (Behrens et al., 2007). Volatility is a measure of stability of statistical parameters (e.g., mean or variance) in a stochastic system. In reward paradigms, volatility refers to the stability of the probability with which rewards are associated with cues (environmental states or actions). This empirical finding raises the question whether ACC also computes higher order statistics (volatility) in addition to the lower order statistics of reward prediction or prediction error. On the one hand, this hypothesis can elegantly account for both behavioral and fMRI findings in humans. On the other hand, the same hypothesis does not seem to have a precise neurophysiological counterpart, because single unit studies have not documented ACC neuronal populations coding for volatility.

In this paper we will show, using the RVPM, that volatility effects can be parsimoniously explained by means of reward prediction and prediction error computation. In the RVPM the average activity of prediction error neurons over a time bin reflects the reward uncertainty during that time bin. Therefore, the average activity of $\delta^+$ and $\delta^-$ neurons in the RVPM is maximal when the reward probability is $.5$ (i.e., when it is maximally uncertain; see Fig. S1 in Supplementary Material). Here we advance the hypothesis that average ACC activation during outcome monitoring (reward period) is mainly driven by the average prediction error evoked by outcomes. This also implies that ACC response to volatility is mainly due to general level of uncertainty, as the predictive links between cues and outcomes change over time in volatile environments, increasing the average activation of the prediction error units (Silvetti et al., 2011). It is useful to point out here a distinction between expected versus unexpected uncertainty (Yu and Dayan, 2005). In the rest of the paper, when we will refer to uncertainty without any further specification, we will mean “expected uncertainty” (i.e., expected reliability of predictive cues) and not to “unexpected uncertainty” (i.e., unexpected changes of reliability of predictive cues) (Yu and Dayan, 2005). The concept of unexpected uncertainty is closely related to volatility. We propose that ACC calculates prediction error and is sensitive to all types of uncertainty. One consequence is that it is active also in volatile environments.

A crucial prediction is that highly uncertain (but not volatile) environments should activate the ACC more than volatile environments. We test this in the current study. In the first part we investigate, by means of RVPM computer simulations, the effects of different levels of uncertainty and volatility on ACC activity. In the second part of the study we tested the predictions derived from the computer simulations by a model-based fMRI experiment, in which we administered to a group of healthy volunteers a task with monetary rewards in environments with different levels of volatility and uncertainty.

2. Methods

2.1. Simulation

The RVPM consisted in three computational modules (Silvetti et al., 2011) (Fig. 1, see also Supplementary Methods). The ACC module receives afferents from units coding for external cues (C1 and C2; representing either stimuli or planned actions) and from a module simulating brainstem dopaminergic afferents (e.g., from ventral tegmental area, VTA) (Williams and Goldman-Rakic, 1998).

We performed 20 simulations, each including three runs of 72 trials (Fig. 2a). In each run we administered to the RVPM trials consisting in cues probabilistically followed by rewards (see also the Supplementary Methods). In the first run we exposed the model to a stationary (Stat) environment, where
two cues were rewarded with constant probabilities (87% and 33% for C1 and C2 respectively). In the second run, we administered a second stationary (Stat2) environment, where the cues were rewarded with constant (i.e., non-volatile) but highly uncertain probabilities (each 60%). Finally, in the third run, we administered a volatile (Vol) environment, where the probability linking cues and rewards started as in the Stat condition, but the probabilities switched halfway (Behrens et al., 2007).

In every trial, the model was required to decide between cue 1 and cue 2, and then a reward was given with a defined probability. For this purpose, we provided the RVPM with a module (Actor) choosing between the two cues. In particular, a Softmax Actor made a choice between one of the two cues on the basis of the expected reward for each cue (V signal). We used a Boltzmann probability distribution having as argument the V value during the last cue period:

\[ p(C_i) = \frac{e^{V_i/\text{Temp}}}{\sum_i e^{V_i/\text{Temp}}} \]  

where \( p(C_i) \) is the probability of selecting the \( i \)th cue, \( \text{Temp} = 4 \) is the temperature parameter, and \( V_i \) is the V response to the last presentation of the \( i \)th cue. We set the temperature parameter to a quite high value; as a result, there was just a small preference for the cue with the higher expectation (55%). This functioned as a convenient approximation for simulating the behavioral results of Behrens et al. (2007), where subjects often chose the cue with the lowest reward probability. Moreover, it ensured an adequate number of selections of both cues. The Actor module was used only in these simulations and not in the simulations producing the regressor signal for the model-based fMRI analysis (see the RVPM regressor in the following section).

### 2.2. fMRI experiment

#### 2.2.1. Experimental procedure

Twenty-one right-handed healthy volunteers (mean age 23, range 19–26, 14 females) gave written informed consent to participate in the study, which was previously approved by the ethical committee of Ghent University. Two participants were excluded from the study because of excessive head motion during scanning. The protocol administered to each subject consisted of three phases, of which only the third
included fMRI scanning. The first phase consisted in setting
the luminance of the stimuli, the second session consisted of
a training session, and the third phase consisted of the scan-
ning session (see also the Supplementary Methods). Subjects
performed a perceptual task with the same Stat, Stat2, and Vol
conditions used in the computer simulations. Each subject
performed each run (Stat, Stat2, Vol) twice for a total of six
runs, or 432 task trials and 96 catch trials, following the experi-
mental timeline shown in Fig. 3a. Trials were charac-
terized by two different colors (red and blue), coding for
reward probability (Knutson et al., 2000). Fig. 3b shows the
sequence of events occurring within each task trial. One of the
two disks appeared some milliseconds before the other two.
The subjects were asked to press a button indicating which
disk appeared first. The delay between the appearance of the
first disk and the other two was adjusted during the training
period, in order to obtain the desired reward probabilities (see
also Behavioral task section in the Supplementary Material).
The reward consisted in 6 Euro-cents, and it was given after
each correct response. Subjects were not informed about the
association between colors and reward probability. Immedi-
ately after the response, an accuracy feedback sign was dis-
played. Each experimental run was performed in one fMRI
run. The mean total scanning duration was 45
48
00. To main-
tain subjects’ performance stable during the Stat2 epochs (the
60% hit rate was sensitive to learning, fatigue, fluctuations of
attention) the program computed a running mean of the
accuracy and continued adjusting the delay to maintain
performance as close as possible to 60%.

2.2.2. RVPM regressor and design matrix
In order to obtain the RVPM time series predicting the ACC
activity, we administered to our model the same sequence of
stimulus–outcome pairs that each subject received (and
produced) during the fMRI experiment (excluding the training
sessions). We considered as RVPM output the sum of the three
ACC units: V, δ+, and δ−. This allowed us to obtain a “personal
simulation of ACC activity for each single volume (i.e., every TR,
2 sec) of each subject (see visual summary in Fig. 4a). As both the
behavioral choices and the response times (RTs) we adminis-
tered to the model were those performed by the subjects, we did
not implement in these simulations an Actor module. In the
same design matrix we included also the regressors relative to
task trials, catch trials and RTs in order to remove their effects as
possible confounds. For a full description of all the regressors we
used in the General Linear Model, we refer to the Methods
section of the Supplementary Material.

2.2.3. Imaging protocol
We acquired fMRI images by a Siemens Trio 3T scanner, with
an 8-channel head coil. Head movements were minimized by
mild restraints and cushioning. Each brain volume was
acquired with 30 slices T2* weighted, voxel resolution
3.5
3.5
3.5 mm3, repetition time 2 sec, time echo 30 msec,
and interslice distance .75 mm. We used SPM8 (http://www.fil.
on.ucl.ac.uk/spm/) and Marsbar toolbox (Brett et al., 2002) for
data processing. The first four volumes of each run were dis-
carded to allow for magnetic field equilibration. After that, all
images were corrected for head movements and slice

---

Fig. 3 – (a) Experiment timeline. We administered three different SEs: stationary (Stat), stationary with high-level of
uncertainty (Stat2) and volatile (Vol), following the same schema of the simulation (Fig. 2a). (b) Events occurring during
a single trial (we arbitrarily depicted a red trial). First a fixation cross appears; the fixation cross is colored; then the actual
task stimulus is shown (three disks); after the response, feedback is given. The catch trial timeline is described in Fig. S3 of
the Supplementary Methods.
acquisition delays, then normalized to the SPM8 EPI template and spatially smoothed by an isotropic Gaussian Kernel with 8 mm full-width half-maximum.

3. Results

3.1. Simulation

The first step of our study consisted in providing a general prediction of ACC behaviour as a function of different statistical environments (SEs). Fig. 1 shows the architecture of the RVPM [see Silvetti et al. (2011) and Supplementary Methods for a more detailed description]. The model learns to associate a value (reward expectation, V unit activity) to each cue, based on the history of past outcomes (prediction error, \( \delta \) units activity). We presented three different environments to the RVPM by manipulating the SE, varying both uncertainty and volatility (Fig. 2a). In the Stat environment, two cues were rewarded with constant probabilities, in the other stationary (Stat2) environment, the cues were rewarded with highly uncertain probabilities, and finally, in the Vol environment, the probability linking cues and rewards switched (volatility) (Behrens et al., 2007).

As shown in Fig. 2b, during the cue period, the RVPM showed the same average activity for all three epochs \( [F(2,57) = .84 \ p = \text{n.s}] \). This is because the average reward rate was the same for each epoch (60%). However, during the reward period, there was a clear main effect of the SE factor \( [F(2,57) = 53.1, \ p < .0001; \text{Fig. 2b}] \). Post hoc analysis revealed
that the RVPM activity was higher during the Vol run than during the Stat run \([t(19) = 3.91, p < .0001]\). The reason is the higher average activation of \(\Delta^+\) and \(\Delta^-\) units during the Vol run. Indeed, after switching the associations between cues and reward rates, the RVPM performed a re-mapping of cue-reward expectations using the prediction errors computed by the \(\Delta^+\) and \(\Delta^-\) units. However, the RVPM showed higher average activity during the Stat2 (highly uncertain but non-volatile run) than during the Vol run \([t(19) = 7.43, p < .0001]\; see also Fig. 2c).

### 3.2. fMRI study

Subjects performed a perceptual task with probabilistic reward in the same SEs used for the computer simulations (Stat, Stat2, and Vol) (Fig. 3a). Trials were characterized by two different cues, coding for reward probability (Fig. 3b). As described in the Methods sections, reward probability was manipulated by varying the difficulty of the perceptual task. The behavioral results showed that our preset error percentages were very well approximated and RTs were faster in the easy (87%) than in the difficult (33%) condition; details are reported in the Supplementary Results section. We administered the behavioral data generated by each single subject to the RVPM, and used the model output as a regressor to determine which voxels were significantly correlated with the RVPM activity (model-based approach, Fig. 4a). Note that the simulations used to generate the predicted BOLD signal were separate from those described in section 3.1. The simulations from section 3.1 provided general predictions on ACC behavior in different SEs (see Methods and Supplementary Methods).

However, based on the simulation from section 3.1, a quantitative prediction is impossible at a single-subject basis. This is why we constructed a new and “personalized” prediction of the ACC BOLD signal for each single subject, based on his/her behavioral data, to analyze the fMRI data.

Fig. 4b shows the voxels correlated to the model (t-contrast Stat+Stat2+Vol), highlighting the fit between the ACC BOLD signal and the RVPM output. On the bottom right we plotted the effect size related to the RVPM regressor as a function of signal and the RVPM output. On the bottom right we plotted the SE, in an region of interest (ROI) (269 voxels) surrounding the effect size related to the RVPM regressor as a function of signal and the RVPM output. On the bottom right we plotted the Stat contrast, and the Vol–Stat contrast (strict conjunction, in which we tested the Conjunction Null hypothesis [Nichols et al., 2005], see Supplementary Methods). This contrast allowed finding the voxels that satisfied all the following conditions. (1) Being correlated with the RVPM regressor. (2) Showing a higher activation for Stat2 than for Vol environment. (3) Showing a higher activation for Vol than for Stat environment. The active cluster consisted of 51 voxels, each significant at \(p < .05\) FWE voxel-wise corrected for the volume of the ACC ROI we built in the main analysis (which was here used as an inclusive mask, 269 voxels). The activation peak was on coordinates 6 16 50 \([t(18) = 3.01]\). To test the possible presence of voxels more active for Vol than for Stat2 condition (i.e., predicted by volatility theory), we computed another contrast, consisting in the linear contrast Vol–Stat2.

### Table 1 – List of activations resulting from the contrast Stat+Stat2+Vol.

<table>
<thead>
<tr>
<th>p(FWE-cor)</th>
<th>Cluster size (vox)</th>
<th>Z</th>
<th>x  y  z (mm)</th>
<th>Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;.0001</td>
<td>2546</td>
<td>6.4</td>
<td>−8 8 54</td>
<td>pre-SMA</td>
</tr>
<tr>
<td>.001</td>
<td>460</td>
<td>5.35</td>
<td>46 2 28</td>
<td>MFG_R</td>
</tr>
<tr>
<td>.022</td>
<td>260</td>
<td>5.12</td>
<td>12 10 −8</td>
<td>Ventral Stri.</td>
</tr>
<tr>
<td>&lt;.0001</td>
<td>742</td>
<td>5</td>
<td>4 −18 −16</td>
<td>VTA</td>
</tr>
<tr>
<td>.001</td>
<td>495</td>
<td>4.93</td>
<td>44 −66 −2</td>
<td>TOJ_R</td>
</tr>
<tr>
<td>&lt;.0001</td>
<td>592</td>
<td>4.57</td>
<td>38 −6 52</td>
<td>PMd_R</td>
</tr>
<tr>
<td>.037</td>
<td>226</td>
<td>4.55</td>
<td>34 24 0</td>
<td>Alns_R</td>
</tr>
</tbody>
</table>

Ventral Stri. = Ventral Striatum, pre-SMA = pre-Supplementary Motor Area, R = right; L = left. Coordinates are in MNI frame of reference. Clusters were obtained with threshold \(p < .001\) voxel-wise uncorrected, and none extent threshold. The first column of the table indicates the \(p\) values cluster-wise FWE corrected.

To remove a possible effect of the average amplitude signal of the RVPM regressor (that, as we showed in the Simulations paragraph, follows the trend Stat2>Vol>Stat), we ran an additional analysis. We normalized each RVPM regressor (mean = 0, SD = 1), eliminating any effect due to differences in average signal amplitude across conditions. Fig. 5 displays the t-map showing the voxels active for the main RVPM effect, the Stat2–Vol contrast, and the Vol–Stat contrast [strict conjunction, in which we tested the Conjunction Null hypothesis (Nichols et al., 2005), see Supplementary Methods]. This contrast allowed finding the voxels that satisfied all the following conditions. (1) Being correlated with the RVPM regressor. (2) Showing a higher activation for Stat2 than for Vol environment. (3) Showing a higher activation for Vol than for Stat environment. The active cluster consisted of 51 voxels, each significant at \(p < .05\) FWE voxel-wise corrected for the volume of the ACC ROI we built in the main analysis (which was here used as an inclusive mask, 269 voxels). The activation peak was on coordinates 6 16 50 \([t(18) = 3.01]\). To test the possible presence of voxels more active for Vol than for Stat2 condition (i.e., predicted by volatility theory), we computed another contrast, consisting in the linear contrast Vol–Stat2.
This contrast revealed no voxels, within the ACC ROI, exceeding the uncorrected threshold of \( p < .05 \).

Finally, in order to exclude that Stat2 environment accidentally generated an amount of volatility greater than Vol environment (e.g., for local fluctuations of reward rates), we compared the reward rates variances of each SE and ran also a supplementary behavioral study. Both analyses confirmed that the Vol condition was indeed more volatile than the two stationary conditions (Stat and Stat2; see Supplementary Results section).

4. Discussion

We proposed a biologically plausible model of ACC, based on the hypothesis that one of the core functions of ACC is to compute reward predictions and prediction errors (Rushworth and Behrens, 2008). The results presented in this paper support a unified theory of ACC function (Silvetti et al., 2011). This theory holds that computation of reward prediction and prediction error is responsible for most of the experimental findings on ACC function, including those that were previously explained by error detection, error likelihood estimation, conflict monitoring, and volatility. Starting from this framework, our computer simulations and fMRI experiment suggested the ACC response to volatility (Behrens et al., 2007) can be parsimoniously explained by the sensitivity of this area to prediction errors, providing a neurophysiological and computational rationale for the effect. Moreover, our fMRI study also confirmed a crucial prediction of the RVPM, namely, that highly uncertain but non-volatile environments can lead to higher ACC activation than volatile environments. This suggests that the average over-time response of this area is mainly driven by general uncertainty (i.e., by both expected and unexpected uncertainty).

A recent theoretical paper described an ACC model based on similar neurocomputational assumptions (Alexander and Brown, 2011). This model, like the RVPM, also provides an explanation of volatility in terms of prediction error activity. The two models also exhibit a few differences; for instance, the RVPM computes expectations and prediction errors specifically for rewards, whereas Alexander and Brown’s model computes expectations and prediction errors more generally on environmental outcomes. Regardless of the differences between the two models, the convergence of independent theoretical results about the role of prediction error signals in volatile environments corroborates the robustness of the results of the current study. Finally, it is worth noting one relevant difference between our experimental approach and that used in some previous studies on the ACC. We investigated ACC processing of expected value and prediction error for stimuli (cue colors, i.e., environmental states), whereas the ACC has been often investigated in relation to choice selection (Behrens et al., 2007; Kennerley et al., 2011) (i.e., actions). Both theoretical considerations (Actor–Critic framework, see below) and experimental findings (Amiez et al., 2006) indicate that stimulus and action evaluation are probably processed in an analogous way by the ACC, and the results of our study (correlation between RVPM activity and biological ACC activity) corroborate this point of view.

4.1. ACC, prediction error and volatility

These findings expand upon the pioneering observations by Behrens et al. (2007) but, at the same time, introduce some new observations that deserve full discussion. In order to isolate the influence of volatility on ACC activation, Behrens et al. de-confounded the effect of prediction errors by using the signal generated by a Rescorla–Wagner model (1972). This model provided positive and negative responses for positive and negative prediction errors. Here we adopted a different modeling approach, in which the RVPM coding of prediction errors is always positive, and provides an output more similar to the neuronal discharge rate signal (Matsumoto et al., 2007), producing a closer approximation to the biological BOLD response. Moreover, in order to demonstrate that the volatility effect was not due to prediction errors, Behrens et al. documented also an absence of linear correlation between the absolute value of prediction errors and volatility. However, this finding is not in conflict with the fact that ACC responds more strongly to uncertain environments, as the amount of uncertainty is not coded by single values of prediction errors, but rather by the average prediction error (i.e., by the average activity of \( \delta \) units over time). For instance, during the Stat period (87–33% condition) the RVPM showed higher maximal prediction error signals (due to highly unexpected violation of reward predictions) than during the Stat2 period (60–60% condition), but the average activity of \( \delta \) units was higher in the Stat2 period.

4.2. Prediction error, noradrenalin and learning rate

In the same fMRI study, Behrens et al. also provided evidence that the volatility estimation determined the subjects’ learning rate. We acknowledge that modeling the learning rate modulation is beyond the scope of the RVPM, as it has a constant learning rate parameter. However, a dynamic learning rate parameter would not have changed our results, given that it would have led to a faster adaptation of the RVPM to the switches in reward contingencies during the volatile period, and thus to lower average activation of the \( \delta \) units, further increasing the gap between the Vol and the Stat2 condition. Moreover, using the RVPM framework, we can advance a hypothesis on learning rate modulation. Indeed, Yu and Dayan (2005) proposed that the neuromodulator noradrenalin (originating from the brainstem nucleus locus coeruleus, LC) signals unexpected uncertainty. We suggest that learning rate modulation arises from top–down influences from ACC to LC. Being so densely connected to different cortical areas, ACC would be ideally suited to provide LC a high-level summary of cortical processing, and it indeed has projections to LC (Aston-Jones and Cohen, 2005). We propose that LC activation, with its widespread cortical projection, sets learning rates in different cortical areas, including the ACC itself (Berridge and Waterhouse, 2003; Verguts and Notebaert, 2009; Jepma and Nieuwenhuis, 2011). The explicit computational formulation of this hypothesis and its eventual experimental test will be the focus of future research.
4.3. **The Actor–Critic framework**

In the framework of RL, a Critic is a system deputized to evaluate stimuli and possible actions in terms of expected reward, while the Actor is a system that selects actions based on the evaluations provided by the Critic. According to the RVPM, the role of the ACC would be that of a Critic module in the Actor–Critic framework (see also Supplementary Material), creating maps between cues (e.g., external stimuli, actions) and values indicating their value for the organism (Rushworth and Behrens, 2008). Consistent with this, ACC receives input from (high-level) motor areas, which code for actions, and also from the posterior parietal cortex (Devinsky et al., 1995), which can be considered as input coding for external stimuli. In mammals, it is reasonable to identify the Dorsolateral Prefrontal Cortex (DLPFC), basal ganglia and high-level motor areas (like the caudal cingulate zone, CCZ) as the Actor. Indeed these areas receive afferents from ACC (Devinsky et al., 1995), allowing a tight interaction between the two components. Besides ACC, a number of other areas showed activity correlated with the RVPM (Table 1). Some of them are typically involved in RL, such as VTA and ventral striatum (D’Ardenne et al., 2008). The others (namely the right mid frontal gyrus – MFG, the dorsal premotor – PMd and the TOJ) are typically not considered to be involved in reward processing. The right MFG has also been shown to code for prediction errors (Matsumoto et al., 2007) but with a less marked distinction between positive and negative prediction errors (thus, encoding general surprise). Hence, we propose right MFG may encode prediction errors in a similar way in humans, coding mainly for surprise of incoming events (Doricchi et al., 2010). The PMd activation may be interpreted as a reward prediction error signal afferent from ACC (Critic) to the PMd neural populations deputized to select actions (Actor). Finally, with respect to the TOJ, a previous fMRI study found prediction error activity in an area overlapping our TOJ activation (Summerfield and Koechlin, 2008). In the latter study the TOJ activity was related to prediction errors on expectations about the features of incoming visual stimuli, and not on reward expectations. This may also be relevant in our paradigm, because regardless of the reward, the colored cues predicted the incoming visual stimulus given by the feedback shape (Fig. 3b). Thus, the TOJ activity could be interpreted as deriving from prediction errors related to the expectations of the feedback shape.

It is worth mentioning that the AIns was also among the reward-related areas we found. Recently, it has been proposed that the AIns has a complementary role to ACC (Preuschoff et al., 2008): whereas ACC is involved in reward processing and represents both reward and reward prediction errors, the AIns would represent risk and risk prediction errors. The insula is also strongly involved in representing feelings and predictions of feelings in others (Singer et al., 2009). These are not necessarily contradictory findings. In the Actor–Critic framework, the Critic can be multidimensional, representing different aspects of cue valuation. Neurophysiological and lesion studies provided indeed evidences that different cortical areas perform partially segregated roles of Critic, such as reward prediction (ACC, orbitofrontal cortex, OFC) (Schoenbaum et al., 2009; Rushworth and Behrens, 2008) or risks (AIns) (Preuschoff et al., 2008) and costs evaluation (ACC, OFC) (Rushworth and Behrens, 2008). More broadly, the ubiquitous finding of ACC and AIns activation in cognitive tasks is congruent with the fact that cognition and emotion are not separate faculties: in an Actor–Critic system, cognition and emotion are two related components for the purpose of adaptive control (Dolan, 2002; Verguts and Notebaert, 2009).

**Acknowledgments**

MS and TV were supported by Ghent University GOA grant BOF08/GOA/011. We acknowledge the support of Ghent University Multidisciplinary Research Platform “The integrative neuroscience of behavioural control”. RS is supported by the Fund for Scientific Research (Flanders Belgium). All the authors declare having no conflict of interest. The authors thank Fabrizio Doricchi, Marcel Brass, Emiliano Macaluso and Anand Ramamoorthy for useful comments on this work.

**Supplementary material**

Supplementary data related to this article can be found online at doi:10.1016/j.cortex.2012.05.008.

**References**


Silvetti M, Seurinck R, and Verguts T. Value and prediction error in the medial frontal cortex: Integrating the single-unit and systems levels of analysis. Frontiers in Human Neuroscience, 5(75) 2011.


