HAL Archives Ouvertes - France
Author Manuscript
Accepted for publication in a peer reviewed journal.

Published in final edited form as:
Cogn Affect Behav Neurosci. 2007 December ; 7(4): 327–336.

Expectations, gains, and losses in the anterior cingulate cortex

Jérôme Sallet*, **René Quilodran**, **Marie Rothé**, **Julien Vezoli**, **Jean-Paul Joseph**, and **Emmanuel Procyk**
Institut cellule souche et cerveau INSERM : U846, Université Claude Bernard - Lyon I, FR.

Abstract

The anterior cingulate cortex (ACC) participates in evaluating actions and outcomes. Little is known on how action/reward values are processed in ACC and if the context in which actions are performed influences this processing. Here we report ACC unit activity of monkeys performing two tasks. The first tested whether the encoding of reward values is context-dependant i.e. dependant on the size of the other rewards available in the current block of trials. The second task tested whether unexpected events signaling a change in reward are represented. We show that the context created by a block design (i.e. the context of possible alternative rewards) influences the encoding of reward values, even if no decision or choice is required. ACC activity encodes the relative and not absolute expected reward values. Moreover, cingulate activity signals and evaluates when reward expectations are violated by unexpected stimuli indicating reward gains or losses.

Keywords

Animals; Conditioning; Operant; physiology; Electrodes; Implanted; Electrophysiology; Macaca mulatta; Male; Prefrontal Cortex; physiology; Psychomotor Performance; physiology; Reaction Time; physiology; Reward

INTRODUCTION

Decisions rely on the estimation of the outcome that one will get after a choice is made and an action performed. This estimation, the expected value, is based on the representations of expected and obtained rewards -and more generally outcomes- associated to objects or acts through learning (Lee & Seo, 2007). Value is affected by time of occurrence, probability, frequency, valence, and size of rewards. The ensemble of alternative outcomes available in the current environment can also have an effect on expected values.

The processing of rewards involves a so-called reward system, a network of brain structures including the orbitofrontal cortex (OFC), the ventral striatum, and the mesencephalic dopaminergic system and the anterior cingulate cortex (ACC) (Schultz, 2000). OFC activity is modulated by expected outcomes and reflects the relative values assigned by the animals to alternative choices (Padoa-Schioppa & Assad, 2006). Encoding of reward values by dopaminergic neurons is adaptive and relative to the potential outcomes (Tobler et al., 2005). Within the reward system the ACC has a particular position. Its role most likely consists in building outcome-action associations that serve to elaborate with experience the expected value of particular decisions (Kennerley et al., 2006; Rushworth et al., 2004). There is evidence that the ACC is involved in representing expected values (Amiez et al., 2006). Thus one can also expect relative coding of reward values in the ACC.

* Correspondence should be adressed to: Jérôme Sallet sallet@lyon.inserm.fr.

HAL-AO Author Manuscript
HAL-AO Author Manuscript
HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

ACC function also encompasses voluntary selection of behavior and positive and negative outcome evaluations (Amiez et al., 2005; Shima & Tanji, 1998; Walton et al., 2004). Several theories are proposed to explain ACC activation in humans during tasks that involve active performance monitoring (Holroyd & Coles, 2002; Kerns et al., 2004; Ridderinkhof et al., 2004; Rushworth et al., 2004). Signals from ACC have been interpreted in terms of detection of conflict in processing competitive representations, e.g. alternative action plans (Botvinick et al., 2004). These ACC signals could also possibly reflect detection of events at odds with expectations, especially in terms of rewards (Gehring & Willoughby, 2002; Holroyd & Coles, 2002; Rushworth et al., 2004). Indeed, ACC feedback-related activity varies with reward prediction errors during learning (Amiez et al., 2005; Holroyd & Coles, 2002; Matsumoto et al., 2007).

The present experiments were designed to further investigate the representations of expected values. We report recordings in the ACC of monkeys performing two tasks: the scaled-reward task, and the expect task that involved neither decision between alternative choices nor learning of particular reward values. The first task was used to evaluate whether modulations of ACC activity by values would be influenced by context, i.e. by the other rewards available within a block of trials and by the status of reward amount (fixed vs probabilistic). The second task was designed to search for ACC activity related to the evaluation of events signaling a change in outcome, at odds (with gain or loss) with the current expectation.

METHODS

Behavior

Housing, surgical, electrophysiological and histological procedures were carried out according to the European Community Council Directive (1986), the Ministère de l'Agriculture et de la Forêt, Commission nationale de l'expérimentation animale, and the Direction Départementale des Services Vétérinaires (Lyon, France). Each animal was seated in a primate chair within arm's reach of a tangent touch-screen (Microtouch System) coupled to a TV monitor. In the front panel of the chair, an arm-projection window was opened, allowing the monkey to touch the screen with one hand. A computer recorded the position and accuracy of each touch. It also controlled the presentation via the monitor of visual stimuli (colour shapes), which served as light-targets (CORTEX software, NIMH Laboratory of Neuropsychology, Bethesda, Maryland). Eye movements were monitored using an Iscan infrared system (Iscan Inc. USA).

Two male rhesus monkeys were trained in the Scaled-Reward task and Expect task (fig. 1). In each task the stimuli were fixed from one day to another. They were presented in the centre of the screen, were 4×4cm in dimension, and were composed of superimposed shapes of different colours. In both tasks each trial was self initiated by the monkey on touching a starting position (SP) on the screen. The animal had then to hold this position. After a short delay a visual cue appeared at the centre of the screen for 500ms. After a waiting delay (from 1000 to 2500ms in Scaled-Reward task and from 800 to 1700ms in the Expect task) a GO signal appeared. The animal had to touch the stimulus given at the GO-signal. A reward juice was delivered 600ms after the touch.

Scaled-Reward task

The Scaled-Reward task was designed to measure the effect of reward size expectation on ACC activity and to test whether expected reward sizes were encoded in a linear or relative fashion by the firing rates. Moreover different scales of reward were presented in different blocks to test the effect of context (i.e. of the other potential rewards in a block). Two scales of four stimuli-reward couples were used. In each scale, the extremes were associated with fixed

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

reward, and the mediums with probabilistic rewards. Reward values for each stimulus were calculated as follows: Value = Size (ml) × Probability of the reward.

In the lower scale (Scale 1) the stimuli were named A, B, C, and D. In the larger scale (Scale 2) the stimuli were named A′, B′, C′, and D′. Reward value associated with each stimulus is reported in figure 1.

The different values were in Scale 1: A: 0.26 = 0.26ml × 1; B: 0.31 = (0.26 × 0.7) + (0.42 × 0.3); C: 0.37 = (0.26 × 0.3) + (0.42× 0.7); and D: 0.42 = 0.42 × 1 and in Scale 2: A′: 0.42 = 0.42× 1; B′: 0.53 = (0.42 × 0.7) + (0.77× 0.3); C′: 0.66 = (0.42 × 0.3) + (0.77× 0.7); D′: 0.77 = 0.77× 1

Importantly, in order to test the effect of context, the reward value of A′ was equal to the reward for D, with the difference that A′ is the minimum reward of Scale 2 and D is the maximum reward of scale 1.

Scales were presented in blocks of trials that were usually interspaced by rest or by a block of trials of the Expect task. Thus the time-interval between tests could vary from 1 to 15 minutes. Each scale was presented once or twice depending on the stability of recordings.

Expect Task

The expect task was designed to test whether ACC generates specific signals at the onset of events that indicate a discrepancy between the expected reward and the reward to be obtained.

In three conditions, three stimuli were associated with small (0.21ml), medium (0.48ml), or large (0.96ml) rewards. In about 75 % of Small and Large trials and 100% of Medium trials, the GO signal was identical to the cue, thus confirming the expected reward. In 25% of Small and 25% of Large trials the GO signal did not match the cue. After a cue indicating a small reward the item referring to a large reward appeared and indicated MORE than expected. After a cue indicating a large reward the item referring to a small reward appeared indicating LESS than expected. There were 5 different types of trials: Large-large, medium-medium, small-small, large-LESS, and small-MORE. The procedure for mixing conditions was made by providing in the random selection additional trials for incongruent conditions. Thus, at the beginning of trials the cues Small, Medium, and Large had an average probability to appear of 0.37, 0.27, and 0.36 respectively (computed from 1687 trials).

Behavioral Measures

Monkeys were required to respond within defined time limits for reaction time and movement times. We tested various time limits from 300 to 600ms delay for reaction time and 300 to 600 ms for the movement time. These constraints were adapted online to allow the animal to perform at an appropriate performance level - to give sufficient numbers of rewards. Reaction time for arm movements (from GO to start-position release) and the number of errors were measured. Errors of several types (No start, NoGo, late Response, Early response) were analyzed separately.

Recordings

Monkeys were implanted with a head-restraining device, and an atlas-guided (Paxinos) craniotomy was done to expose an aperture over the prefrontal cortex. A recording chamber was implanted with its centre placed at stereotaxic anterior level +31. A vertical chamber implanted vertically on the midline was used for monkey P. A lateral chamber positioned at 32 degree from the vertical plane was used for monkey M. Neuronal activity was recorded using epoxy-coated tungsten electrodes (1–4 MOhm at 1 kHz; FHC inc, USA). One to four

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

microelectrodes were placed in stainless steel guide tubes and independently advanced into the cortex through a set of micromotors (Alpha-Omega Engineering, Israel). Neuronal activity was sampled at 13kHz resolution. Electrode tracks and recording locations were verified using anatomical MRI for monkey P and histology for monkey M. Recording sites covered an area extending over about 6mm (anterior to posterior), in the dorsal bank of the anterior cingulate sulcus, at stereotaxic antero-posterior levels superior to A+30, and at depths superior to 4.5mm from cortical surface. This corresponds to a region recorded in previous reports and in which error-related activity have been observed (Amiez et al., 2006; Ito et al., 2003; Procyk et al., 2000). This part of the anterior cingulate cortex lies at the same anterior level as the SEF, and includes part and goes anterior to the rostral cingulate motor area (CMAr) as evaluated from previous publications (Ito et al., 2003; Shima et al., 1991).

Single unit activity was identified using online spike sorting (MSD, AlphaOmega). Activity of single neurons was analyzed with respect to different events and outcomes (conditions) by using averaged PSTH and trial by trial spike counts. Statistical evaluations of neural activity were done using average spike frequency measured in separated time epochs delineated in each trial (NeuroExplorer, Nex Technology, USA; MatLab – The MathWorks Inc.). PSTHs had a binning of 0.01s and were Boxcar averaged. Average population activity was smoothed using a Lowess fitting (precision 0.02; Statistica, Statsoft Inc.).

Analyses

For the *Scaled-Reward task*, we selected 7 epochs: **SP:** onset starting position to 300ms after touch of the position; **Cue:** 500 msec period following cue onset; **Delay:** the first 500ms of the delay; **Go:** from the Go signal to the starting position release; **Movement:** from the release lever to target touch; **Post Touch:** from target touch to reward delivery, and **Reward:** 600msec following reward delivery. A baseline epoch was defined as the last 500ms of the ITI. Expected reward size effects were tested using a 2-ways ANOVA with epoch x reward as factors ($p<0.05$).

One issue is whether the neural activity encodes expected reward values monotonically. To address this problem, we first selected one particular epoch per cell. We used a selectivity index based on a nonparametric one-way analysis of variance (Kruskal-Wallis test, factor REWARD SIZE). The analysis extracted a Chi2 value. We used it as a measure of the variance explained by REWARD SIZE. The epoch in each cell with the largest Chi2 value was selected for further analyses. A normalization was then adopted that (i) transformed activity in trial A (A′) and trial D (D′) to '0' and '100', respectively, and (ii) linearly adjusted the value observed in the other trials to this new scale. Let X be the average activity measured in one epoch; we normalized X by:

$$X \rightarrow 100 \times (X - A)/(D - A)$$

where A and D represent the average activity for the same epoch in trials A and D. With this transformation, the activity in trial A and in trial D in particular became equal to 0 and 100 respectively, independently of whether activity in trial D was initially larger or smaller than in trial A (fig. 5). The normalization was applied using activity for trials A and D as reference for neurons having significantly different activity in trials A and D. Activity in A′ and D′ were used as references for neurons having significantly different activity in trials A′ and D′.

For the *Expect task*, change in neural activity (neural response) was tested for being event-related. A change in activity was considered to be significant if the average activity measured within the 500ms post-event (post-Cue or post-GO), in at least one condition, exceeded more than twofold the standard deviation of the pre-event average activity (taken during the window −600/−200 ms preceding the event alignment time), and remained above this threshold for more than six 0.01s bins. Average activity was measured by detecting the peak in average

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

activity (all conditions combined) in the 500ms following event times and computing the average activity in a −160ms to +160ms window around the peak time. The average activity was standardized to the mean and standard deviation of the baseline activity taken from −600 to −200ms before event onset. Resulting data were expressed in number of sigma (SD) of the baseline.

Average population activity on single units showing significant effects was constructed to illustrate the statistical results. This format was also used to observe the dynamic of the average activity in the different subpopulations.

RESULTS

1 Scaled-Reward task. Expected values and ACC activity

Behavior—Errors in monkey P (fig. 2A) were primarily due to No-Go responses (no arm-movement towards the target) and to a small proportion of no-start (trial not initialized). Error rates tended to be smaller for large rewards within each scale. RTs were statistically the same in all trials. In monkey M (fig. 2B), rates of errors, due to No-Go trials, were smaller than in monkey P. They also tended to be smaller for large rewards. RTs did not vary within scales, only between scales. A global modulation of error rate by reward size expectation in scale 1 (assessed by the ANOVA; $p<0.05$) was observed for both monkeys. Furthermore, for both monkeys, we observed a significant difference in performance (error rate or RTs) between D trials and A′ trials indicating a context effect (fig. 2 A–B).

In summary, the performance of the two monkeys was modulated by the expected size of rewards although the context effect (D vs A′) and the block effect on errors were different between the two monkeys.

Neuronal data—This Scaled-Reward task was designed to assess how ACC neurons encode reward values. Our a *priori* hypothesis was that ACC encodes relative reward values in a context dependant manner. Thus the relative coding of rewards in one definite scale should be similar in another scale even if absolute reward values are different.

We first evaluated whether the neural activity recorded for each cell was modulated by the size of expected rewards. Figure 4A shows the activity of a typical example cell. A 2-way ANOVA (EPOCH x REWARD VALUE, threshold $p<0.05$) was performed independently for the two scales. Out of 62 cells, recorded for at least one reward scale (60 cells recorded for scale 1 - the smaller rewards- and 57 for scale 2 -the larger rewards), 24 and 25 showed a main effect of REWARD VALUE and/or an interaction between REWARD VALUE and EPOCH. Thus activity of 40 and 45% of ACC cells was modulated by expected reward values.

To characterize the type of coding (linear, tuned…) performed by the cells we worked on a group of cells (n=23) that fulfilled the following requirements: - they were recorded in both scales, - they revealed an effect of reward value in at least one scale (ANOVA see above), - and their baselines for scale 1 and 2 were not significantly different (Wilcoxon rank sum test, $p<0.05$). We then isolated in these cells the epochs for which the activity was showing the strongest effects of reward value that is, with maximum reward index (see methods). On average maximum indices were found at epochs surrounding arm movements suggesting that the best discrimination of reward values was taking place in association to movement planning and execution (fig. 3).

We plotted, for the 23 cells the average activity measured in the selected epoch. 6 examples are presented in figure 4B. Various relationships between expected reward values and neural activity were found. The preferred expected value (the one for which maximum activity was

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

found) within a scale could be either the smallest (A or A′) the largest (D or D′) or one of the intermediate. The encoding of expected reward sizes was different from cell to cell as would be expected if cells had different tunings.

The experiment was designed specifically to test whether encoding of expected reward values in one definite scale is similar in another scale, even if absolute reward values are different (relative encoding).

We first tested whether Scale and Reward were two interacting factors. A 2-ways ANOVA ($p<0.05$; Scale x Reward-Rank; Reward-Rank: ranks are 1, 2, 3 and 4 for A, B, C and D and for A′, B′, C′, and D′) was performed on the epochs with maximum index. From 14/23 cells showing a Reward-Rank effect (at $p<0.05$) only 4 showed an interaction with the factor Scale. Statistical results are show in figure 4B for the examples. The absence of interaction in 10/14 cells (2/3rd) reveals the cases for which there is no change in relative encoding between the scales. Thus a majority of the analyzed cells showed a relative and adaptive coding of expected rewards.

To further compare the two scales we selected the cells for which the average activity differed between the extreme rewards (A vs D and/or A′ vs D′). A Wilcoxon rank-sum test (at $p<0.05$) performed on epochs presenting the maximum index of selectivity revealed 11 cells with significant differences between A and D and/or between A′ and D′ (one cell showed an effect for both tests). Average population data for the 11 cells are shown in figure 5. The left panel shows the average responses for 6 neurons with significantly different activity for A and D. For each neuron the 4 averages of activity recorded in each scale (8 values) has been normalized so that activity for A is at 0 and activity for D is at 100 (see methods). Data are plotted against absolute reward values on the abscissa. The right panel shows the average for other 6 cells (one cell is also contributing to the graph in the left panel) with significantly different activity for A′ and D′, and normalized using average activity for A′ and D′.

The two plots in figure 5 confirm at the population level that there is no monotonic increase or decrease according to the absolute reward values from one scale to another. Although the differences between rewards in one scale tend to disappear in the other scale, the averages stay in the same range. The graph shows as a comparison the activity that would be expected if it were linearly coding for the relative reward values within a scale (dashed lines).

In summary, 45% of cells were modulated by expected reward sizes. However the subpopulation of cells we analyzed showed that the encoding of expected reward sizes available in one block is relative and not absolute.

2 Expect task. Unexpected gains or losses and ACC activity

Behavior—In monkey P (fig. 2C), errors were due to No-Go responses. The proportion of No-Go trials was larger for small rewards (trials “Small” and “Less”). The error rate was also higher in LESS than MORE conditions. RTs were shorter for small rewards. Note that the proportion of No-Go errors was quite large and created problems to get sufficient data in the LESS condition.

Monkey M (fig. 2D) made fewer No-Go trials. Error rates and RTs were similar for small and large rewards, and for “MORE” and “LESS” conditions. Significant differences were observed in error rates only in Small vs. Medium, Medium vs. Large and Large vs. MORE conditions.

In summary, the effects of reward schedules (Reward size, and MORE vs LESS) were similar in the two monkeys but weaker for monkey M.

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

Neuronal data—95 cells (59 in monkey M and 36 in monkey P) were recorded long enough to study activity at the GO signal in correctly performed congruent and incongruent trials (trials in which the monkey received a reward).

For 42 neurons (44%) the activity significantly increased within the 500ms after the onset of at least one of the 3 cues (> 2SD from pre-Cue; see methods). Of these cue-related activities, 45% (n=19) showed an effect of cue identity/expected reward size (ANOVA $p<0.05$).

Out of the 95 cells, 56 (59%) showed significant GO-related activity, that is a significant increase in the 500ms after the GO signal in at least one of the five conditions (LESS, MORE, Small, Medium, and Large) (> 2SD from pre-GO; see methods).

We were especially interested in whether cells could show distinct activity for MORE and/or LESS conditions after the GO signal. 16 cells had significant (>2SD) *and* maximum average activity at the GO signal for the MORE condition compared to Small, Medium, LESS, and Large. 18 other cells had significant *and* maximum average activity for the LESS condition compared to Small, MORE, Medium and Large. Two examples are shown in figure 6. Figure 6A illustrates the phenomenon for 'MORE' conditions. In the "small-MORE" condition the MORE item (which is identical to the Large item) induces higher activity in this cell compared to all other conditions. There is no particular discharge to the unexpected 'Small'. A similar effect is found for LESS neurons (see a single example in figure 6B). The MORE and LESS effects are observed in the two population activity (fig. 7 A, C). These data reveal that the ACC is signaling the discrepancy between the expected reward and the reward announced at the time of the GO signal, even before any reward is delivered.

Modulation of activity at the Cue and at the different GO signals could possibly reflect a pure effect of probability of occurrence. The MORE item has a probability of 0.25 to appear after a cue indicating a small reward; the 'Large' cue has a probability 0.36 of being presented, and the expected Large GO a probability of 0.75. In this hypothesis the activity should increase or decrease with probabilities. This was not the case for both LESS and MORE cells (fig. 7B and 7D). This suggests that the MORE and LESS signals reflect the unexpected gain or loss in reward.

An important question concerns the specificity of ACC responses, whether they are always coding in one direction (MORE or LESS) or whether there are other activities reacting to both incongruent conditions (MORE and LESS) without specifying gain or loss. No cell was found to have significant activity in *both* MORE and LESS conditions in comparison to Small and Large that would have revealed cells signaling pure incongruent GO signal.

DISCUSSION

Our data show that ACC activity encodes expected reward values and that the coding is relative to the context. In our protocol the context was the ensemble of alternative reward values given in the current block of trials. In addition, ACC unit activity signals the violation of expectations in terms of gain or loss of reward.

Limitations

The effects of reward size expectation or of incongruent conditions on behavioral performances were present, although not strictly identical, for the two animals. It is possible that the online adjustments of time constraints might have slightly influenced the speed and accuracy of responses for each animal. This procedure was used to adapt to each animal and avoid no-response or late-response errors that prevented us to record neuronal activity for sufficient number of trials. At the same time it forced the animal to attend to the GO signal. Other

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

procedures might be considered to avoid behavioral discrepancies between subjects. In any case, the effects of experimental conditions on neuronal activity in the two monkeys were comparable.

The number of task-related neurons used to compare activity in scale 1 and 2 in the Scaled-reward task is low. Criteria on trial numbers also induced important data selection for the Expect task for which there was a limited number of correct trials in the LESS condition. Conservative criteria were defined in order to avoid experimental biases. These criteria increased the relevance of our statistical results.

For both tasks we used behavioral situations in which there was no challenge in terms of response choice. This might have led us to record ACC at its minimal involvement. Indeed several lines of evidence suggest that ACC is much involved in situations where actions are self-selected, outcomes uncertain, and for which high behavioral flexibility is required (Procyk et al., in press; Rushworth et al., 2007; Rushworth et al., 2004). Yet, the recorded task-related activities showed different coherent effects of context and reward size expectations in the two tasks.

Expectations of rewards

Amiez et al. have shown that in a probabilistic learning task where decisions are adapted to maximize rewards, ACC neurons code for the average reward value of the best choice, the task-value (Amiez et al., 2006). Here we show that ACC activity is modulated by expected reward sizes even in a simple task that involved neither decision between several alternatives choices nor learning of action values.

Within-scale near monotonic encoding of reward quantities could be inferred from cells in which response to D is different from response to A, or response to D′ is different from response to A′ (fig. 5). For a sub population of cells, the activity in anticipation of fixed rewards were overall larger or smaller than the activity for intermediate rewards (D and C, D′ and C′). Interpretation of these responses is not straightforward. It may reflect the different reward schedules (certain *vs* probabilistic). A monotonic pattern of value coding in ACC has been observed previously in a choice task (Amiez et al., 2006). A monotonic coding of reward values might be a prerequisite to make optimal decisions in probabilistic environments. Thus it might have been less likely to occur using the present protocol. On the other hand non monotonic coding might reflect the existence of tuning curves that reveal the general mechanism of information processing by local cortical networks. Such mechanism, observed for reward processing in other frontal areas (Padoa-Schioppa & Assad, 2006;Wallis & Miller, 2003), could also apply to the ACC. Further studies will be needed to investigate further the encoding properties of ACC neurons and the respective influences of context and reward schedule.

Importantly our data show that within a particular context (block of trials, range of available rewards) values are encoded relative to each other. This reveals adaptive properties of neural coding in ACC, and suggests that relative encoding of rewards can be adjusted to any learned context. Adaptive coding is not specific to the ACC. Other brain structures (mesencephalic dopaminergic neurons, OFC, and striatum) encode relative reward values of expected or obtained rewards within a scale (Cromwell & Schultz, 2003; Tobler et al., 2005; Tremblay & Schultz, 1999). The adaptation of neural encoding is an important mechanism for the reward system to encode an infinite number of possible reward values in specific or changing environments.

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

Violation of reward expectations

In situations where reward size expectations were violated by an incongruent event (Expect task), the ACC showed activities that were signaling a gain or loss in reward. This adds to our previous report of ACC error-related activities modulated by expectations (Amiez et al., 2005). In human ERP experiments a mid-frontal negavity called medial frontal negativity (MFN) was found to reflect gain or losses (Gehring & Willoughby, 2002). The origin of the MFN is supposed to be in the ACC. We show here the first unit activity data that directly support this hypothesis. Unit recordings confirm that ACC not only detects unexpected events but also evaluates them in terms of gain or loss in reward. At the population level our data reveal that ACC encodes and discriminate both gains and losses (that is discrepancies in terms of rewards). Since the populations coding for gains and losses are segregated within the ACC it is likely that the output of the area concerns different neural populations in target structures and thus possibly induces different responses, for instance in terms of adaptation.

Our experiment shows that this computation can take place in ACC even before any reward is delivered, by comparison of the value associated to a reward-predicting cue with the current expectation. This is comparable to the MFN observed during a “slot machine” task, in which evaluation of outcomes where not contingent of any preceding choice or action (Donkers et al., 2005). Importantly, the Expect task involves a motor response that arises just after the evaluative signals; thus any activity we recorded here can be interpreted in terms of action valuation.

Conclusions

Major theoretical positions regarding ACC function relate to error detection, conflict monitoring, and outcome-action association. They converge on the notion that ACC is detecting and evaluating unexpected events that relate to goal achievement, and is using this evaluation to trigger adaptive mechanisms.

Our data show first that ACC activity encodes expected reward values, gains and losses. Thus our data support a role of ACC in value estimation and computation of prediction errors which are two key elements for reinforcement learning. One important result is that the encoding of values is relative. In other words, reward values are encoded in relation to alternative rewards (the context) and not on an absolute scale. Coding relative values is of primary importance for a structure involved in behavioral adaptation and internally guided decision making.

Acknowledgements

Our thanks go to H. Kennedy for helpful comments and language editing. As part of our research on neurobiology of executive functions these experiments benefited from funds from Inserm, Région Rhônes Alpes, Agence National de la Recherche, Fyssen and NRJ foundations. EP and JPJ are employed by the CNRS. JS is funded by Ministère de I’Education et Recherche and Fondation pour la Recherche Médicale, MR by Ministère de I’Education et Recherche, JV by Région Rhône Alpes, and RQ by Facultad de Medicina Universidad de Valparaiso, projecto MECESUP UVA-106, and Fondation pour la Recherche Médicale.

References

Amiez C, Joseph JP, Procyk E. Anterior cingulate error-related activity is modulated by predicted reward. Eur J Neurosci 2005;21(12):3447–3452. [PubMed: 16026482]

Amiez C, Joseph JP, Procyk E. Reward encoding in the monkey anterior cingulated cortex. Cereb Cortex 2006;16(7):1040–1055. [PubMed: 16207931]

Botvinick MM, Cohen JD, Carter CS. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn Sci 2004;8(12):539–546. [PubMed: 15556023]

Cromwell HC, Schultz W. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. J Neurophysiol 2003;89(5):2823–2838. [PubMed: 12611937]

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

Donkers FC, Nieuwenhuis S, van Boxtel GJ. Mediofrontal negativities in the absence of responding. Brain Res Cogn Brain Res 2005;25(3):777–787. [PubMed: 16249075]

Gehring WJ, Willoughby AR. The medial frontal cortex and the rapid processing of monetary gains and losses. Science 2002;295(5563):2279–2282. [PubMed: 11910116]

Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev 2002;109(4):679–709. [PubMed: 12374324]

Ito S, Stuphorn V, Brown JW, Schall JD. Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science 2003;302(5642):120–122. [PubMed: 14526085]

Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF. Optimal decision making and the anterior cingulate cortex. Nat Neurosci 2006;9(7):940–947. [PubMed: 16783368]

Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS. Anterior cingulate conflict monitoring and adjustments in control. Science 2004;303(5660):1023–1026. [PubMed: 14963333]

Lee D, Seo H. Mechanisms of reinforcement learning and decision making in the primate dorsolateral prefrontal cortex. Ann N YAcad Sci. 2007

Matsumoto M, Matsumoto K, Abe H, Tanaka K. Medial prefrontal cell activity signaling prediction errors of action values. Nat Neurosci 2007;10(5):647–656. [PubMed: 17450137]

Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. Nature 2006;447(7090):223–226. [PubMed: 16633341]

Procyk, E.; Amiez, C.; Quilodran, R.; Joseph, JP. Modulations of prefrontal activity related to cognitive control and performance monitoring. Attention and Performance. 22nd Attention & Performance Meeting: Sensorimotor Foundations of Higher Cognition; Pizay, France. in press

Procyk E, Tanaka YL, Joseph JP. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. Nat Neurosci 2000;3(5):502–508. [PubMed: 10769392]

Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. The role of the medial frontal cortex in cognitive control. Science 2004;306(5695):443–447. [PubMed: 15486290]

Rushworth MF, Behrens TE, Rudebeck PH, Walton ME. Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. Trends Cogn Sci. 2007

Rushworth MF, Walton ME, Kennerley SW, Bannerman DM. Action sets and decisions in the medial frontal cortex. Trends Cogn Sci 2004;8(9):410–417. [PubMed: 15350242]

Schultz W. Multiple reward signals in the brain. Nat Rev Neurosci 2000;1(3):199–207. [PubMed: 11257908]

Shima K, Aya K, Mushiake H, Inase M, Aizawa H, Tanji J. Two movement-related foci in the primate cingulate cortex observed in signal-triggered and self-paced forelimb movements. J Neurophysiol 1991;65(2):188–202. [PubMed: 2016637]

Shima K, Tanji J. Role for cingulate motor area cells in voluntary movement selection based on reward. Science 1998;282(5392):1335–1338. [PubMed: 9812901]

Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. Science 2005;307(5715):1642–1645. [PubMed: 15761155]

Tremblay L, Schultz W. Relative reward preference in primate orbitofrontal cortex. Nature 1999;398 (6729):704–708. [PubMed: 10227292]

Wallis JD, Miller EK. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. Eur J Neurosci 2003;18(7):2069–2081. [PubMed: 14622240]

Walton ME, Devlin JT, Rushworth MF. Interactions between decision making and performance monitoring within prefrontal cortex. Nat Neurosci 2004;7(11):1259–1265. [PubMed: 15494729]

HAL-AO Author Manuscript
HAL-AO Author Manuscript
HAL-AO Author Manuscript

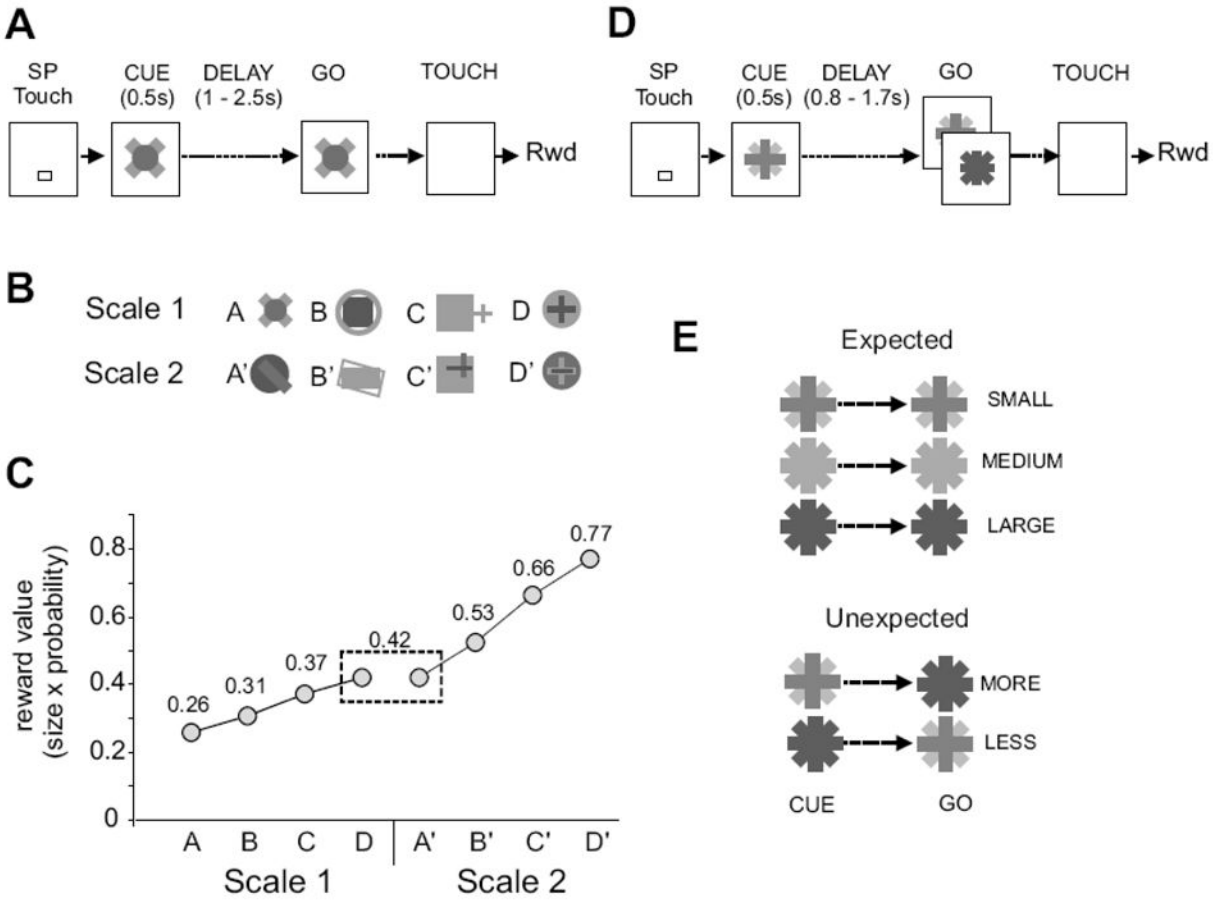

Figure 1. Behavioral tasks. A–C, Scaled-reward task

A. Schematic representation of the scaled-reward task. The animal started a trial by touching the starting position (SP) and holding his touch. Then a Cue switched on for 0.5s. The Cue indicated the amount of reward delivered in case of successful trial. After a variable delay, the item switched on again (Go signal), and the animal could release the starting position and touch the item. 600ms after the touch, the animal received the reward. **B.** Cues used for the 4 reward conditions in each scale. Cues from one scale were presented in block of trials. **C.** Graphic showing the reward value (size × probability) associated with each reward condition in each scale. **D–E, Expect task. D.** Schematic representation of the Expect task. The task is similar to the reward-scaled task described in A. The major difference is that in some trials the GO item is different from the Cue (in about 25% of trials). **E.** Representation of the Cue/Go pairs used in this study. In 3 conditions, the GO is identical to the Cue and the expected reward is delivered. In two conditions, the Go item is different from the Cue and the reward to be obtained is larger (MORE) or smaller (LESS) than the expected reward.

HAL-AO Author Manuscript
HAL-AO Author Manuscript
HAL-AO Author Manuscript

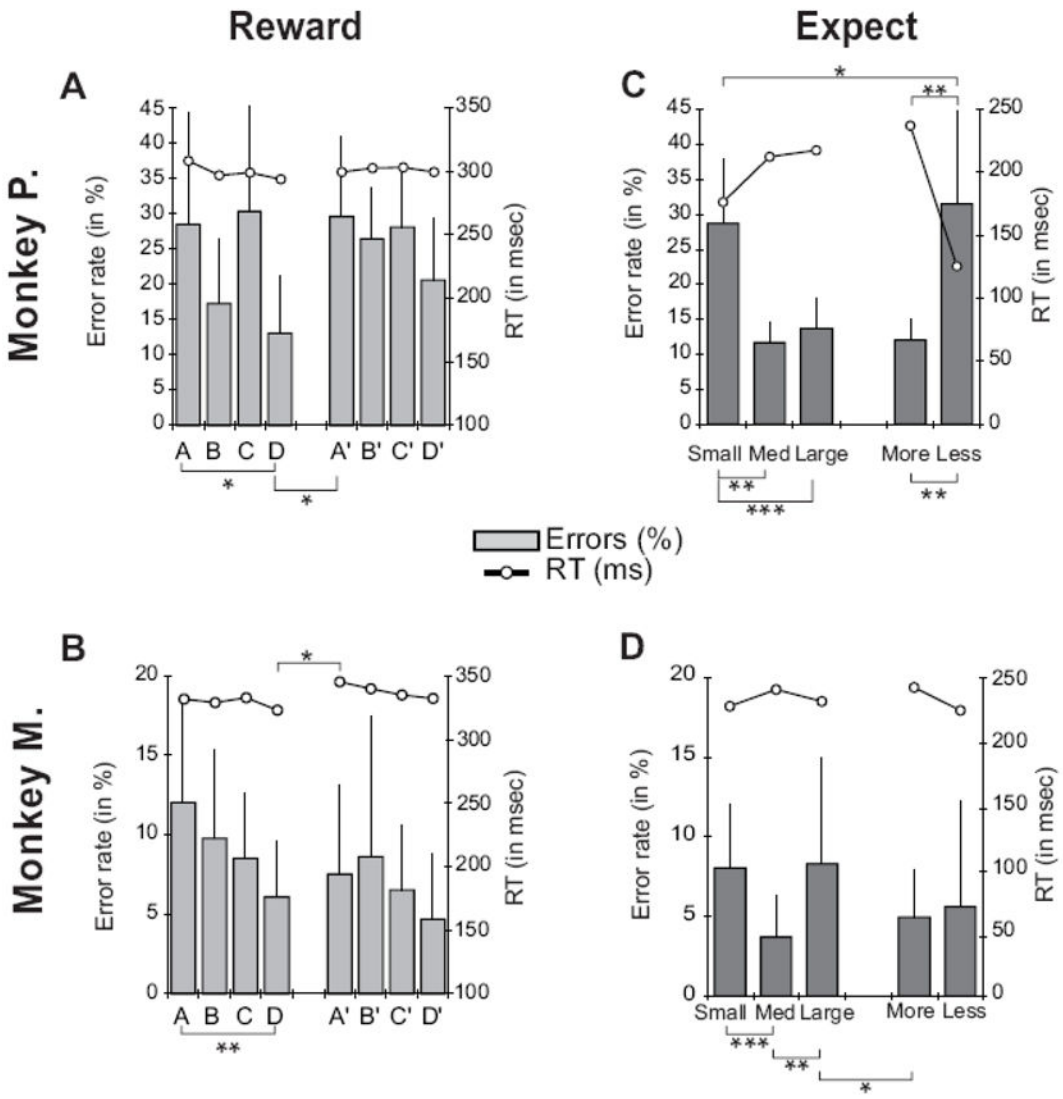

Figure 2. Behavioral data for monkeys P and M in the Scaled-Reward (A–B) and in the Expect tasks (C–D)

Error rates (Bars) and Reaction times (RT; plots). One-way ANOVA at * p<0.05; ** p<0.01; ***: p<0.001. Significance of tests are shown on top for RT and below histograms for Errors.

HAL-AO Author Manuscript

HAL-AO Author Manuscript

HAL-AO Author Manuscript

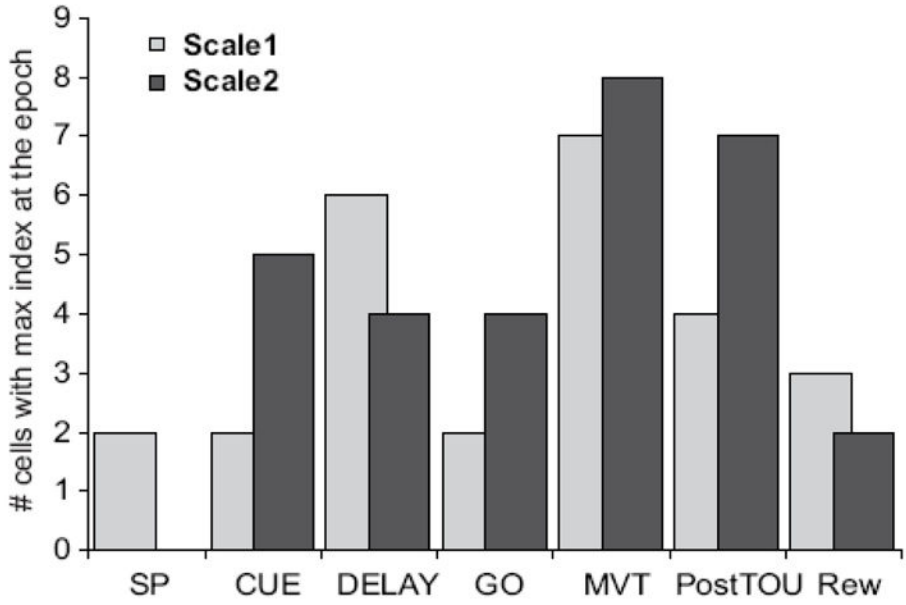

Figure 3. Scaled-reward task

Distribution of epochs with maximum reward index (revealed by a Chi2). Data are given for Scale 1 and 2 separately.

HAL-AO Author Manuscript
HAL-AO Author Manuscript
HAL-AO Author Manuscript

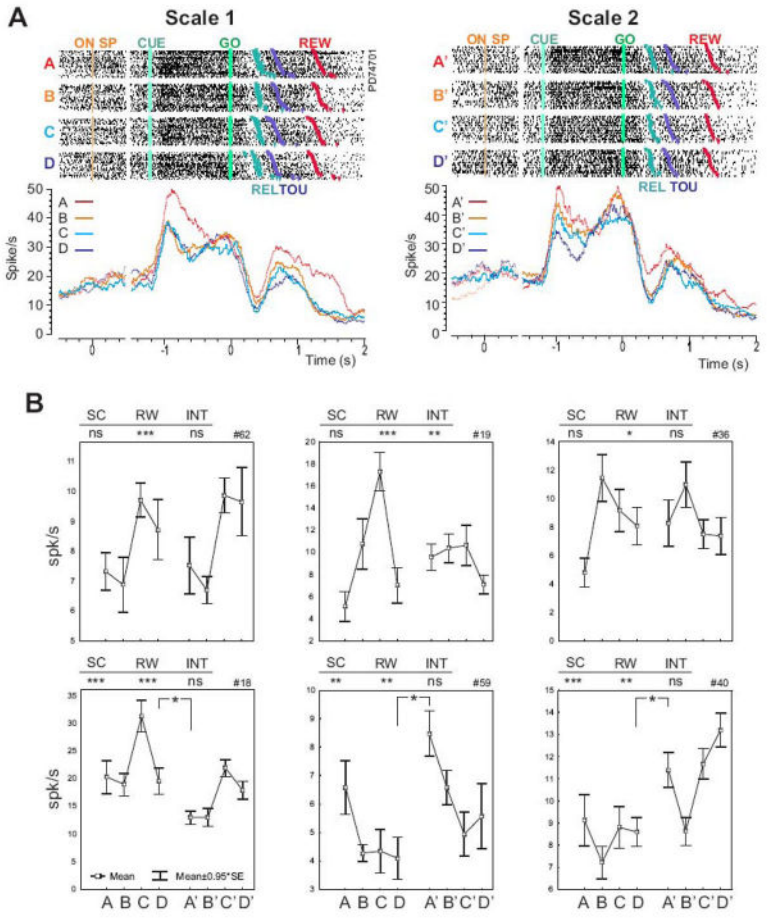

Figure 4. Neuronal data in the scaled-reward task

In A, activity of one neuron that varies with reward values. Rasters and histograms are presented for scale 1 (left) and scale 2 (right) for each reward value (colors). Activity is aligned on the onset of the starting position (ON SP), the cue onset (CUE) and on the Go Signal (GO). Colored triangles in rasters indicate the release of the starting position (REL), the touch of the target (TOU) and the reward delivery (REW). In B, average activity for 6 individual cells. Each panel represent the 8 averages corresponding to the reward values of the 2 scales (A, B, C, D and A ′, B′, C′, D′ respectively). Statistical difference between average activity in D and A′ are indicated inside the plot (Wilcoxon; *: p<0.05). Above each panel the results from an ANOVA [SCALE (df=1) × RW (df=3)] testing the effect of scale and relative reward values. The significance of each factor and of interactions are indicated (*: p<0.05; **:p<0.01; ***:p<0.001). Note the similarity between the distributions of averages found in the two scales.

HAL-AO Author Manuscript
HAL-AO Author Manuscript
HAL-AO Author Manuscript

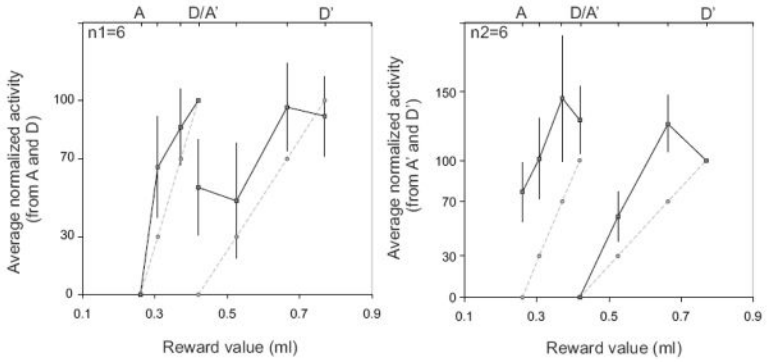

Figure 5.

Average Normalized activity in the two scales for 11 cells with different activity in A and D (left; n1=6) and/or in A′ and D′ (right; n2=6). In each panel the average activity for each reward value is normalized according to the values in A and D (left) or A′ and D′ (right). The abscissa is in ml and the corresponding reward value names are displayed on the top. D and A′ have the same value. The graph shows as a comparison the activity that would be expected if it were linearly coding for the relative reward values within a scale (dashed lines).

HAL-AO Author Manuscript
HAL-AO Author Manuscript
HAL-AO Author Manuscript

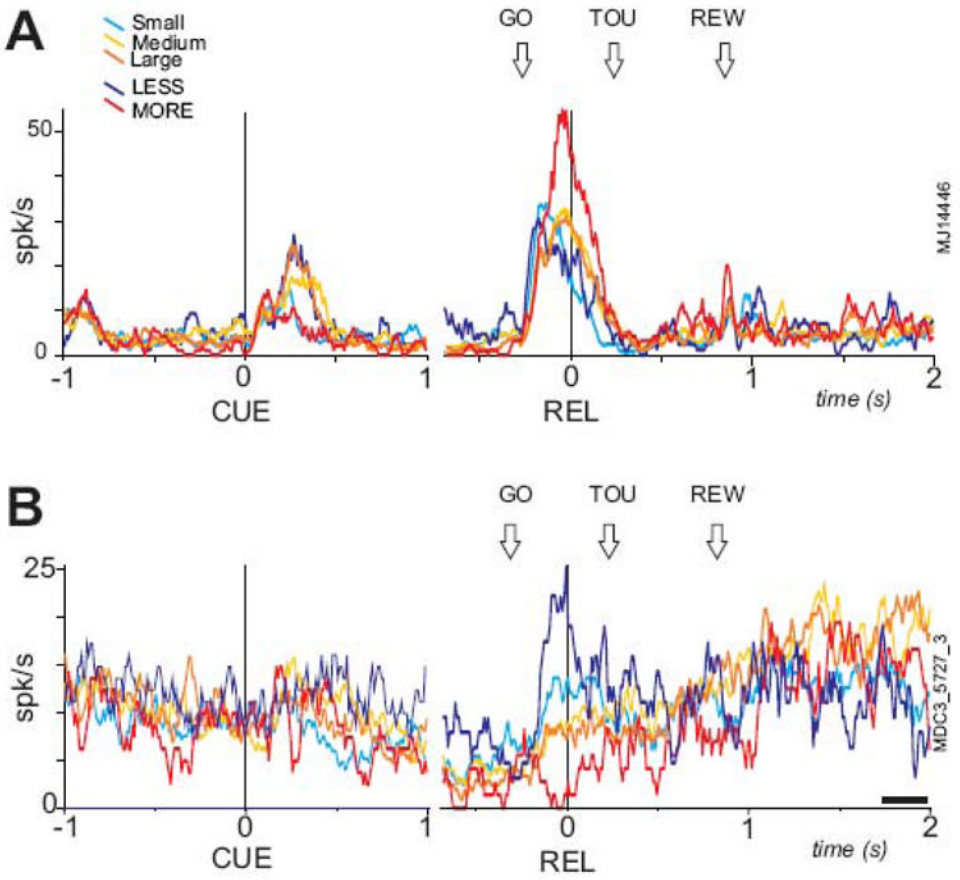

Figure 6. Examples of neurons in the Expect task

Two examples of unit activities recorded in the 5 conditions. The histograms are aligned on the cue (CUE) on the left side of the figure, and on the release of the starting position (REL), on the right side. Colored lines correspond to the different conditions (Large: dark orange dark, Medium: light orange, Small: blue, MORE: red, and LESS: dark blue). The cell in **A** shows an increase of activity at the cue that is statistically larger for cues announcing larger rewards than small rewards (Small and MORE). The cell is more active at the GO signal for the MORE condition. The cell in **B** presents a maximum activity at the GO signal for LESS, and lower activity for the other conditions.

HAL-AO Author Manuscript
HAL-AO Author Manuscript
HAL-AO Author Manuscript

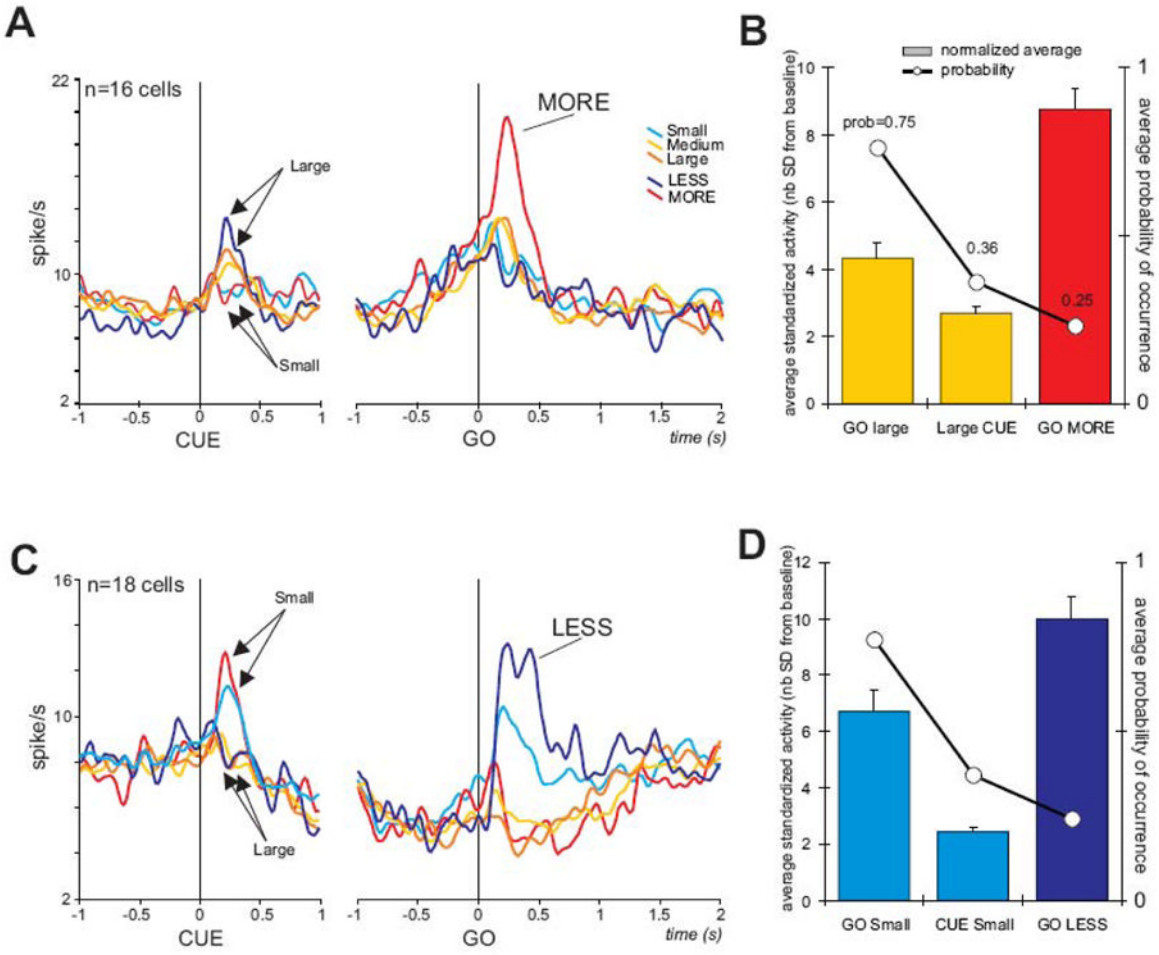

Figure 7. Population activity for two groups of cells in the Expect task

A. Average population activity for cells showing significant activity and maximum average at the GO signal in "Small-MORE" trials. Average activity obtained for each condition is aligned on CUE and GO separately. Colors indicate conditions as in figure 5. Note that there is congruence between reward size preference at the cue and at the GO signal, and that there is a strong magnification effect combined to a reward size effect at the GO signal when the GO item indicates an unexpected reward size. **B**. Bar graph of average standardized activity ([activity-mean baseline]/SD baseline; see methods) for the different events ranked according to their average predictability in trials. The predictability (probability) is indicated on the superimposed plot. The figure shows that response to events is not following the probability. **C**. Average activity for cells showing significant activity and maximum average at the GO signal in "Large-LESS" trials. **D**. Same as C but for the LESS population.